Algorithms seminar, 1993–1994

Bruno Salvy

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE

Algorithms seminar, 1993-1994

Bruno SALVY, éditeur scientifique

N ° 2381

Octobre 1994

PROGRAMME 2

ISSN 0249-6399

ALGORITHMS SEMINAR,

1993–1994

Bruno Salvy

(Editor)

Abstract

These seminar notes represent the proceedings of a seminar devoted to the analysis of algorithms and related topics. The subjects covered include combinatorial models and random generation, symbolic computation, asymptotic analysis, average-case analysis of algorithms and data structures, and computational number theory.

SÉMINAIRE ALGORITHMES,

1993–1994

Abstract

Ces notes de séminaires représentent les actes, pour la plupart en anglais, d'un séminaire consacré à l'analyse d'algorithmes et aux domaines connexes. Les thèmes abordés comprennent : modèles combinatoires, génération aléatoire, calcul formel, analyse asymptotique, analyse en moyenne d'algorithmes et de structures de données, ainsi que de la théorie algorithmique des nombres.

ALGORITHMS SEMINAR
1993–1994

Bruno Salvy[1]
(Editor)

Abstract

These seminar notes represent the proceedings of a seminar devoted to the analysis of algorithms and related topics. The subjects covered include combinatorial models and random generation, symbolic computation, asymptotic analysis, average-case analysis of algorithms and data structures, and computational number theory.

This is the third of our series of seminar proceedings. The previous ones have appeared as INRIA Research Reports numbers 1779 and 2130. These summaries, usually written by a reporter from the audience, form the content of these proceedings.

The primary goal of this seminar is to cover the major methods of the average-case analysis of algorithms and data structures. Neighbouring topics of study are combinatorics, symbolic computation and asymptotic analysis.

Several articles deal with classical combinatorial objects, their enumeration according to various parameters, or their random generation, useful for simulations and empirical studies.

Computer algebra plays an increasingly important rôle in this area. It provides a collection of tools that permit to analyse complex models of combinatorics and the analysis of algorithms; at the same time, it inspires the quest for developing ever more systematic solutions to the analysis of well characterized classes of problems. In this vein, the notes contain several recent developments regarding the automatic resolution of differential equations.

Asymptotic methods include singularity analysis, the saddle point method, Rice's method or Mellin transform techniques.

The 28 articles included in this book represent snapshots of current research in these areas. A tentative organization of their contents is given below.

PART I. COMBINATORIAL MODELS

In addition to its own traditions rooted in mathematics, the study of *combinatorial models* arises naturally in the process of analyzing algorithms that often involve classical combinatorial structures like strings, trees, graphs, permutations.

Combinatorial interpretation of continued fractions and Padé approximants are described in [**1**]. In [**2**], a general technique for random generation of unlabelled combinatorial objects is described. An introduction to the q-calculus is given in [**3**]. Various types of words and of parameters on words are the subject of the next three articles. The enumeration of overlap-free words over a two-letters alphabet is considered in [**4**]. Descent in words are related to several q-analogues of classical quantities in [**5**] and similar parameters for words with repetition are studied in [**6**].

[1]This work was supported in part by the ESPRIT III Basic Research Action Programme of the E.C. under contract ALCOM II (#7141).

PART II. SYMBOLIC COMPUTATION

Most exactly solvable models of combinatorics and analysis of algorithms rest on a suitable algebra of *generating functions*. Once this has been recognized, an important goal is to find decision procedures for classes of generating functions. Computer algebra systems provide a way of testing and implementing the methods, and the problem of optimizing the corresponding procedures often represents a non trivial problem of symbolic computation.

An important class of generating functions is formed by solutions of differential equations. Any kind of solution to these differential equations can be used to help the analysis. In the linear case, very nice algorithms exist to find Liouvillian solutions [7]. Other results in a related direction are presented in [8]. It is sometimes useful to get numerical estimates of the singularities, and a new algorithm in this context is given in [9]. Fast algorithms for computer algebra are crucial to extend computer algebra to real-life problem. Such algorithms form the subject of [10].

PART III. ASYMPTOTIC ANALYSIS

Asymptotic analysis is an essential ingredient in the interpretation of quantitative results supplied by the resolution of combinatorial models.

An important class of problems involves recovering the asymptotic form of the coefficients of a function from asymptotic properties of the function itself. In the most general case, one may have to resort to real analysis and compute subtle bounds before applying any theorem [11]. Fortunately, large classes of expressions can be attacked by general methods. Alternating sums arising from the calculus of finite differences are candidate to the application of Rice's method [12]. Harmonic sums can be attacked by Mellin transforms [13]. Other problems may require an asymptotic knowledge of iterates of analytic functions, a short introduction to simple cases [14] shows how hairy things can be.

PART IV. ANALYSIS OF ALGORITHMS AND DATA STRUCTURES

This part deals with the analysis of algorithms and data structures. Trees have been recognized by various authors as the single most important structure in computer science. Not unnaturally, several analyses found here are devoted to tree structures and their generalizations.

The first four articles of this part deal with asymptotic probability distributions. General theorems that enable to predict the limit distribution of many parameters of combinatorial structures are the subject of [**15,16,17,18**].

The rest of this part is devoted to the analyses of specific algorithms and data-structures. String-searching is studied in [**19**]. Factorization of polynomials is analyzed in [**20**]. Quadtrees of any dimension are dealt with in [**21**]. A function due to Ramanujan which arises in several analyses of algorithms is discussed in [**22**]. The last two articles [**23,24**] deal with sizes of relations in relational databases.

[**15**] Special Limit Distributions for Combinatorial Structures. *Michèle Soria*
[**16**] Limiting Distributions in Product Schemas. *Michèle Soria*
[**17**] Limit Theorems for Combinatorial Structures. *Hsien-Kuei Hwang*
[**18**] *Factorisatio Numerorum*, Combinatorial Constructs and Gaussian Laws. *Hsien-Kuei Hwang*
[**19**] Average-Case Analysis of String-Searching. *Mireille Régnier*
[**20**] Random Polynomials and Factorization Algorithms. *Xavier Gourdon*
[**21**] The Cost Structure of Quadtrees. *Bruno Salvy*
[**22**] Ramanujan's Q-function and Computer Science Applications. *Helmut Prodinger*
[**23**] Sizes of Relations: a Dynamic Analysis. *Danièle Gardy*
[**24**] Data Base Parameters: Equijoin and Semijoin. *Guy Louchard*

PART V. MISCELLANY

This part contains an introduction to elliptic functions and modular forms [**25**] and their applications in modern algorithms in computational number theory [**26**]. Decidability results for some ODEs are considered in [**27**] and a review of genetic algorithms is given in [**28**].

[**25**] Elliptic Functions and Modular Forms. *François Morain*
[**26**] Implementation of the Schoof-Atkin-Elkies Algorithm. *François Morain*
[**27**] PCD Systems and Their Algorithmic Properties. *Eugène Asarine*
[**28**] État de l'art des algorithmes génétiques. *Evelyne Lutton*

Acknowledgements. The lectures summarized here emanate from a seminar attended by a community of researchers in the analysis of algorithms, in the Algorithms Project at INRIA (Ph. Flajolet, F. Morain and B. Salvy are the organizers) and in the greater Paris area—especially École Polytechnique (J.-M. Steyaert), University of Paris Sud at Orsay (D. Gouyou-Beauchamps) and LITP (M. Soria).

The editor expresses his gratitude to the various persons who supported actively this joint enterprise, most notably: Philippe Dumas for his careful rereading, Xavier Gourdon, Dominique Gouyou-Beauchamps and Michèle Soria. Thanks are due also to the speakers and to the authors of summaries. Many of them have come from far away to attend one seminar and nicely accepted to write the summary.

We are especially indebted to Virginie Collette for permanently keeping the wheels in motion.

The Editor
B. SALVY

Part 1

Combinatorial Models

Combinatorial Interpretations of Continued Fractions

Emmanuel Roblet
LaBRI, Bordeaux

June 13, 1994

[summary by Philippe Flajolet]

Abstract

This seminar describes historical motivations and a combinatorial setting for continued fraction expansions of formal power series. By general theorems, universal continued fractions are generating functions of lattice paths in the plane. This can be used either to solve counting problems in terms of continued fractions or to develop a combinatorial approach to continued fraction identities. Roblet's presentation develops this theory applied not only to standard continued fractions (the so-called J- and S-fractions), but also to Padé approximants, T-fractions, and 2-point approximants.

1. Euler

The story begins with an initial frustration of Euler —*"What to do with a divergent series of a differential equation?"*— turned into a brilliant intuition —*"Expand into continued fractions!"*. Euler's intuition was to be later brought to fruition by Stieltjes; it is historically the starting point of orthogonal polynomials, a domain that has since blossomed with rich applications to approximation theory. The talk concerns provides a historical background and then proceeds to discuss the formal aspects of numerous identities relating power series, continued fractions, orthogonal polynomials and classical combinatorial structures.

Euler starts with the purely divergent series

$$Y(z) = \sum_{n=0}^{\infty} n! z^n. \tag{1}$$

One reason for investigating this series is that it directly arises when one attempts to solve the differential equation

$$z\frac{d}{dz}(zy(z)) - y(z) + 1 = 0, \tag{2}$$

by indeterminate coefficients. What if Nature was confronting us with such an equation? Which "information" is concealed in a series like $Y(z)$?

Euler's idea is to try expanding (1) into a *continued fraction*, and he then discovers a remarkable

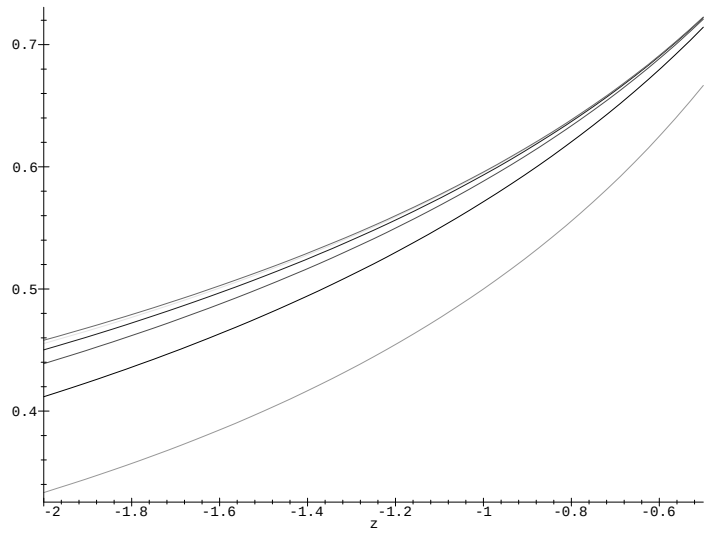

FIGURE 1. The first five convergents of Euler's continued fraction, starting with $P_1/Q_1 = 1/(1-z)$, converge quickly to a well-defined limit.

pattern,

$$\sum_{n=0}^{\infty} n! z^n = \cfrac{1}{1 - 1\cdot z - \cfrac{1^2 \cdot z^2}{1 - 3\cdot z - \cfrac{2^2 \cdot z^2}{1 - 5\cdot z - \cfrac{3^2 \cdot z^2}{\ddots}}}}, \tag{3}$$

involving simply odd numbers and squares.

The format of (3) may seem strange at first sight but it is not. Given a formal power series $f(z) = \sum_{n=0}^{\infty} f_n z^n$, define its integral and fractional parts by

$$\lfloor f \rfloor = f_0 + f_1 z, \qquad \{f\} = f_2 + f_3 z^2 + \cdots,$$

so that

$$f = \lfloor f \rfloor + z^2 \{f\}.$$

The continued fraction (3) is then obtained by applying a simple variant of the usual continued fraction algorithm to Y, very much like in the well-known arithmetic case (take inverses and iterate on fractional parts).

Returning to the original problem, one may wonder what to do with (3) in the context of standard analytic solutions to the original equation (2). One naturally looks at the sequence of *convergents*,

$$\frac{0}{1}, \quad \frac{1}{1-z}, \quad \frac{1-3z}{1-4z+2z^2}, \quad \cdots \quad \frac{P_h(z)}{Q_h(z)}, \cdots$$

Consider them for instance for $z \in [-2, -\frac{1}{2}]$. The miracle is that they do appear to converge to a well defined function, $Y_1(z)$, as shown in Figure 1.

It turns out that $Y_1(z) = \lim_{h\to+\infty} \frac{P_h(z)}{Q_h(z)}$ is an actual *analytic* solution to (2) in the complex plane slit along $[0, +\infty]$. This fact is not too hard to establish in our particular case since everything is fairly explicit: the P and Q polynomials satisfy, like any polynomials arising from convergents, a

linear recurrence of order 2; as the coefficients are simple, the recurrence is solvable and we just obtain the (reciprocal polynomials) of the Laguerre polynomials and their associates.

About a century later, Stieltjes was to prove that this is a general phenomenon and he gave general conditions under which a (possibly divergent) series is numerically approximated by the particular *rational approximants*, P_h/Q_h. (At this point the connection with differential equations is lost as one addresses a much more general problem, namely summation of divergent series.)

In addition, as is well-known, the Laguerre polynomials are orthogonal with respect to the scalar product

$$\langle u, v\rangle = \int_0^\infty u(x)v(x)e^{-x}\,dx,$$

with moments

$$n! = \int_0^\infty x^n e^{-x}\,dx.$$

Hence we have the representation of the original series as a *Stieltjes transform* of e^{-x}:

$$Y_2(z) = \int_0^\infty \frac{1}{1-zx} e^{-x}\,dx.$$

Here, the well defined $Y_2(z)$ is "consistent" with $Y_1(z)$ in the sense that $Y_1(z) = Y_2(z)$ for z in the slit plane.

Once more, this is a general fact: the reciprocals of the denominator polynomials Q_h are always orthogonal (at least formally, and under Stieltjes' conditions analytically as well) with respect to a scalar product

$$\langle u, v\rangle = \int_0^\infty u(x)v(x)\,d\mu(x),$$

as shown by a simple algebraic computation. The full continued fraction is then the Stieltjes transform of the orthogonality measure $d\mu(x)$,

$$\int \frac{1}{1-zx}\,d\mu(x),$$

which may be viewed as a continuous version of a partial fraction decomposition of the original function (or its associated continued fraction).

Though orthogonal polynomials are also useful for enumerations, the rest of this presentation concentrates on enumerative uses of continued fractions. At any rate, continued fractions are the historical source of the theory of orthogonal polynomials, starting from Euler's example.

Note. Classical bibliographical sources for the theory alluded to here are the books by Perron [8] and Wall [14]. Chihara's book [1] is a useful introduction to orthogonal polynomials. Stieltjes' Collected Papers [11] have been recently reprinted with an insightful summary by Van Assche of Stieltjes' contributions.

2. Jacobi and Stieltjes

Euler's idea generalizes. Given a series

$$J(z) = 1 + \sum_{n=1}^{\infty} j_n z^n,$$

the continued fraction algorithm almost surely succeeds and delivers a continued fraction expansion of the form

$$J(z) = \cfrac{1}{1 - \kappa_0 \cdot z - \cfrac{\lambda_1 \cdot z^2}{1 - \kappa_1 \cdot z - \cfrac{\lambda_2 \cdot z^2}{1 - \kappa_2 \cdot z - \cfrac{\lambda_3 \cdot z^2}{\ddots}}}}, \tag{4}$$

called a J-fraction (for Jacobi). A variant of the J-fraction is the S-fraction (for Stieltjes)

$$S(z) = \cfrac{1}{1 - \cfrac{\lambda_1 \cdot z}{1 - \cfrac{\lambda_2 \cdot z}{1 - \cfrac{\lambda_3 \cdot z}{\ddots}}}}. \tag{5}$$

The two are closely related and it is easily recognized that

$$\text{S-fraction}[f(z)] \underset{z^2 \mapsto z}{=} \text{J-fraction}[f(z^2)].$$

Again, this notion is general as a series almost surely admits an S-fraction expansion. (Only rational functions and functions whose Taylor expansion "resembles" a rational function do not admit of such continued fraction expansions.)

3. Continued fractions and special functions

Stieltjes also determined the continued fraction for the ordinary generating function of the Bell numbers $\{B_n\}$ in the form

$$\sum_{n=0}^{\infty} B_n z^n = \cfrac{1}{1 - 1 \cdot z - \cfrac{1 \cdot z^2}{1 - 3 \cdot z - \cfrac{2 \cdot z^2}{1 - 5 \cdot z - \cfrac{3 \cdot z^2}{\ddots}}}}, \tag{6}$$

while a continued fraction expansion of Gauß related to hypergeometric functions implies

$$\sum_{n=0}^{\infty} (2n-1)!! z^n = \cfrac{1}{1 - \cfrac{1 \cdot z^2}{1 - \cfrac{2 \cdot z^2}{1 - \cfrac{3 \cdot z^2}{\ddots}}}}, \tag{7}$$

where $(2n-1)!! = 1 \cdot 3 \cdot 5 \cdots (2n-1)$.

At this stage, it is of interest to note that (3), (6), (7) are Laplace transforms of certain exponential generating functions that all have a simple form, namely

$$\frac{1}{1-z}, \quad e^{e^z - 1}, \quad e^{z^2/2}.$$

In effect the simplest proof of these continued fractions is by means of a theorem of Stieltjes and Rogers that relates an addition formula for an exponential generating function $\phi(z)$ to the continued fraction expansion of the corresponding ordinary generating function $f(z)$. Consider therefore a Laplace pair,

$$f(z) = \sum_{n=0}^{\infty} f_n z^n \quad \text{and} \quad \phi(z) = \sum_{n=0}^{\infty} f_n \frac{z^n}{n!}.$$

The Stieltjes-Rogers theorem asserts that any addition formula for ϕ in the form

$$\phi(x+y) = \sum_{k=0}^{\infty} \phi_k(x)\phi_k(y) \quad \text{where} \quad \phi_k(x) = O(x^k)$$

is simply converted into a J-fraction expansion of $f(z)$. For instance for $\phi(z) = \sec(z) = 1/\cos(z)$, the addition formula reads

$$\phi(x+y) = \frac{1}{\cos(x+y)} = \sum_{k=0}^{\infty} \frac{\sin^k x}{\cos^{k+1} x} \frac{\sin^k y}{\cos^{k+1} y},$$

which corresponds to the expansion

$$f(z) = \sum_{n=0}^{\infty} S_n z^n = \cfrac{1}{1 - \cfrac{1^2 \cdot z^2}{1 - \cfrac{2^2 \cdot z^2}{1 - \cfrac{3^2 \cdot z^2}{\ddots}}}}, \tag{8}$$

with $S_n = n![z^n]\sec(z)$ a secant number.

4. Combinatorics of continued fractions

Continued fractions like (3), (6), (7), (8) with such regular patterns cannot leave a combinatorialist indifferent as the Taylor coefficients count unconstrained permutations, set partitions, involutions, and alternating permutations respectively. At the same time, the continued fractions have coefficients given by simple integral laws.

A basic theorem due to Touchard [13], Good [3], Lenard [6], Szekeres [12], Jackson [5], Flajolet [2], and Read [9] expresses a general connection between J- and S-fractions on the one hand, lattice walks in $\mathbb{Z} \times \mathbb{Z}$ on the other hand. It is no accident that it arises naturally in calculations of random walk probabilities [3], and it is strongly connected with combinatorial configurations equivalent to lattice paths, like chord systems [9, 13]. The original inspiration for [2], is the dynamic analysis of data structures ("histories" introduced by Françon), while Jackson [5] and Lenard [6] draw their motivations from certain one-dimensional models of statistical physics (the Ising model and electrostatics, respectively). Szekeres [12], on the other hand, started from scattered observations of Ramanujan regarding general continued fractions. (Apart from [2], a description of the basic theory may be found in the book by Jackson and Goulden [4, Ch. V].)

Consider the path on the integer lattice $\mathbb{Z} \times \mathbb{Z}$ made of three types of steps

$$\textit{ascents } \tbinom{1}{1},\ \textit{descents } \tbinom{1}{-1},\ \text{and } \textit{levels } \tbinom{1}{0},$$

that start at the origin, finish at altitude 0, and are constrained to stay in the upper-right quarter plane. Each walk can be encoded multiplicatively by a (noncommutative) monomial defined by associating

$$1,\ \lambda_j,\ \kappa_j,$$

to a step of type ascent, descent with starting altitude equal to j, level with altitude j, respectively. The walk polynomial W_n is the sum of all monomial encodings of all walks made of n steps.

THEOREM 1 (TGLSJFR). *The universal J-fraction is the generating function of the set of walk polynomials*

$$\cfrac{1}{1-\kappa_0\cdot z-\cfrac{\lambda_1\cdot z^2}{1-\kappa_1\cdot z-\cfrac{\lambda_2\cdot z^2}{1-\kappa_2\cdot z-\cfrac{\lambda_3\cdot z^2}{\ddots}}}}=\sum_{n=0}^{\infty}W_n z^n\equiv 1+\kappa_0 z+(\kappa_0+\lambda_1)z^2+\cdots .$$

This theorem was in effect used to establish combinatorially all the continued fraction expansions (3), (6), (7), (8), by appealing in an essential way to combinatorial bijections of Françon and Viennot. Conversely, any enumeration problem that can be reduced to a weighted lattice path counting leads to a continued fraction expression for the corresponding ordinary generating function. From there, the whole arsenal of special function identities can then be employed.

An illustrative example is the counting of "coin fountains" by Odlyzko and Wilf. An n-fountain is an arrangement of coins in rows such that the first row has no gaps and each coin in a higher row touches exactly two coins in the next lower row (see Example 10.7 of [7]). The contour of an n-fountain clearly resembles a legal lattice path, and from the TGLSJFR theorem (by adapting the weights), one gets for the ordinary generating function of fountains the representation

$$F(q)=\sum_{n=0}^{\infty}F_n q^n=\cfrac{1}{1-\cfrac{q}{1-\cfrac{q^2}{1-\cfrac{q^3}{\ddots}}}}.$$

The continued fraction is expressible in terms of q-exponentials and an analysis of its dominant polar singularity furnishes the very precise asymptotic formula:

$$F_n\sim C\cdot A^n \qquad \text{with} \qquad C\simeq 0.31236,\quad A\simeq 1.73566.$$

A comparable process has led to the solution of several counting problems like shared communication networks (in terms of Hermite polynomials, by Lagarias and Odlyzko), urn models of the Ehrenfest type (by Goulden and Jackson), correlated Gaussian variables (by Odlyzko *et al.*), non-overlapping partitions (by Flajolet and Schott), etc.

In any case, a continued fraction representation is usually the centre of a rich cluster of identities involving a specific class of orthogonal polynomials, Hankel matrices, and Stieltjes matrices.

5. Some special combinatorial fractions

Because of space constraints, we can only allude to some of the other combinatorial continued fractions discussed by Roblet in his lecture. See [10] for details.

Permutations. The continued fraction

$$\cfrac{1}{1 - f\cdot z - \cfrac{tpq[1]_q^2\cdot z^2}{1 - q(fq+(f+r)[1]_q)\cdot z - \cfrac{tpq^3[2]_q^2\cdot z^2}{1 - q^2(fq^2+(f+r)[2]_q)\cdot z - \cfrac{tpq^5[3]_q^2\cdot z^2}{\ddots}}}}$$

is the ordinary generating function of permutations counted according to inversions (marked by q), and the cyclic structure: fixed points (marked by f), peaks of cycles [*i.e.*, $\sigma^{-1}(i) < i > \sigma(i)$] (marked by p), troughs [*i.e.*, $\sigma^{-1}(i) > i < \sigma(i)$] (marked by t), double rises (marked by r), and double falls (marked by f). This development relates to a bijection of Biane and it provides a 6-parameter statistic on permutations. (There, as usual $[k]_q = 1 + q + \cdots + q^{k-1}$.)

Another bijection of Roblet and Viennot yields a trivariate statistics on permutations by means of the continued fraction

$$\cfrac{1}{1 - (ab-[a;q]_1)\cdot z - \cfrac{[a;q]_1\cdot z}{1 - (bq-[a;q]_2)\cdot z - \cfrac{[a;q]_2\cdot z}{1 - (bq^2-[a;q]_3)\cdot z - \cfrac{[a;q]_3\cdot z}{\ddots}}}},$$

with a marking right-to-left minima and b marking left-to-right maxima, with q again for inversions. (There, $[a;q] = a + q + q^2 + \cdots + q^{k-1}$.) Notice that the above fraction is not a J-fraction, as the numerators involve z instead of z^2. It is known as a T-fraction (for Thron). Obtaining such continued fractions require changing the notion of legal lattice path, and it is one of the major contributions of [10] to develop a systematic theory of such fractions.

Polyominoes. Parallelogram polyominoes can be encoded by various types of lattice paths (Delest, Viennot, Fédou). Roblet gives a T-fraction representation for the joint statistics of height, width, area, and perimeter.

6. Padé approximants and T-fractions

Roblet's thesis also proposes a general theory of continued fraction expansions that differ from the basic J- or S-type. We have already encountered the examples of T-fractions.

An important class of rational approximants is that of Padé fractions. A general combinatorial approach for them is given in [10]. This necessitates appreciably modifying the notion of legal paths. A valuable result is then to explain combinatorially the possible degeneracies in the Padé table, while simply interpreting the basic algebraic identities of the theory.

From his combinatorial interpretations, Roblet is able to deduce systematic (and novel) algorithms for the expansion of a power series into various types of continued fractions. The design is based on a generalization of the Stieltjes matrix that in the classical case is related to the ϕ_k in the Stieltjes-Rogers addition formula and combinatorially to path terminating at an altitude different from 0. The algorithms so obtained have many desirable features: they are "incremental" (a useful feature for computer algebra applications) and of low complexity, consuming space $O(n)$ and time $O(n^2)$. Again, we have to refer to Roblet's thesis for details.

Bibliography

[1] Chihara (T. S.). – *An Introduction to Orthogonal Polynomials*. – Gordon and Breach, New York, 1978.

[2] Flajolet (P.). – Combinatorial aspects of continued fractions. *Discrete Mathematics*, vol. 32, 1980, pp. 125–161.

[3] Good (I. J.). – Random motion and analytic continued fractions. *Proceedings of the Cambridge Philosophical Society*, vol. 54, 1958, pp. 43–47.

[4] Goulden (Ian P.) and Jackson (David M.). – *Combinatorial Enumeration*. – John Wiley, New York, 1983.

[5] Jackson (D.). – Some results on product-weighted lead codes. *Journal of Combinatorial Theory, Series A*, vol. 25, 1978, pp. 181–187.

[6] Lenard (A.). – Exact statistical mechanics of a one-dimensional system with Coulomb forces. *The Journal of Mathematical Physics*, vol. 2, n° 5, September 1961, pp. 682–693.

[7] Odlyzko (A. M.). – Asymptotic enumeration methods. – Preprint, March 1993. To appear as a chapter in the *Handbook of Combinatorics*, R. Graham, M. Grötschel and L. Lovász, ed.

[8] Perron (Oskar). – *Die Lehre von der Kettenbrüchen*. – Teubner, 1954, vol. 2.

[9] Read (Ronald C.). – The chord intersection problem. *Annals of the New York Academy of Sciences*, vol. 319, May 1979, pp. 444–454.

[10] Roblet (Emmanuel). – *Une interprétation combinatoire des approximants de Padé*. – PhD Thesis, Université de Bordeaux I, 1994.

[11] Stieltjes (Thomas Jan). – *Œuvres Complètes*. – Springer Verlag, 1993. Edited by Gerrit van Dijk.

[12] Szekeres (G.). – A combinatorial interpretation of Ramanujan's continued fraction. *Canadian Mathematical Bulletin*, vol. 11, n° 3, 1968, pp. 405–408.

[13] Touchard (Jacques). – Sur un problème de configuration et sur les fractions continues. *Canadian Journal of Mathematics*, vol. 4, 1952, pp. 2–25.

[14] Wall (H. S.). – *Analytic Theory of Continued Fractions*. – Chelsea Publishing Company, 1948.

Random Generation of Unlabelled Combinatorial Structures

Paul Zimmermann
INRIA Lorraine

October 25, 1993

[summary by Eithne Murray]

Abstract

A systematic method for generating unlabelled combinatorial structures uniformly at random is presented. It applies to structures that are decomposable, i.e., formally specifiable by grammars involving unions, products, sequences, multisets and cycles. All such structures of size n can be generated uniformly by either sequential algorithms of worst-case arithmetic complexity $\mathcal{O}(n^2)$ or by "boustrophedonic" algorithms of worst-case complexity $\mathcal{O}(n \log n)$. The random generation procedures are derived automatically from a high level description of the combinatorial structures. An implementation of this system in the computer algebra system Maple is briefly described.

1. Introduction

Presented here is a systematic method for generating unlabelled combinatorial structures at random. Given a grammar-like specification of a class $\mathcal{C}$ of such structures, there is a method to automatically produce procedures for the random generation of objects in $\mathcal{C}$ of a fixed size n. The analysis of the labelled case has already been done [1]. The unlabelled case is more complicated because of the symmetries in unlabelled objects.

The specification language consists of union, cartesian product, sequence, multiset and cycle, as well as two basic initial objects: the structure ϵ of size 0, and Z, an atomic node of size 1. Our method can be used on any *decomposable* class of objects that can be specified by using these constructions iteratively or recursively.

Typical examples of decomposable classes are integer partitions, necklaces, unlabelled trees and forests, random mapping patterns, term trees under associative and commutative laws, and derivation trees of all context-free languages. As an example of a specification consider the class $\mathcal{H}$ of unlabelled hierarchies, that are defined as non-plane trees in which the degrees of the internal nodes are always at least 2. We define $\mathcal{H}$ with the specification language by

$$\mathcal{H} = Z + \texttt{multiset}(\mathcal{H}, \text{card} \geq 2).$$

(Unlabelled hierarchies are relevant to statistical classification theory.)

2. Counting

If we know the number of objects c_n in a decomposable class $\mathcal{C}$ of size n, we can generate an element of size n uniformly at random. Looking at the ordinary generating functions that correspond to each construction in the specification language, we get the following theorem.

THEOREM 1 (FOLK THEOREM OF COMBINATORIAL ANALYSIS). *Given a specification Σ for a class C, a set of equations for the corresponding generating functions is obtained automatically by the following translation rules:*

$$\begin{cases} C = A + B & \Longrightarrow & C(z) = A(z) + B(z) \\ C = A \times B & \Longrightarrow & C(z) = A(z) \cdot B(z) \\ C = \texttt{sequence}(A) & \Longrightarrow & C(z) = \dfrac{1}{1 - A(z)} \\ C = \texttt{multiset}(A) & \Longrightarrow & C(z) = \exp\left(\displaystyle\sum_{k=1}^{\infty} \frac{1}{k} A(z^k)\right) \\ C = \texttt{cycle}(A) & \Longrightarrow & C(z) = \displaystyle\sum_{k=1}^{\infty} \frac{\varphi(k)}{k} \log \frac{1}{1 - A(z^k)} \end{cases}$$

where $\varphi(k)$ denotes the Euler totient function.

Thus, to count the number of objects of size n of a structure given by a grammar specification, we can use its corresponding enumerating generating function that is built up from the grammar according to the rules of the theorem.

3. Standard Specifications

We transform each grammar into a kind of Chomsky Normal form. Thus, each union and product will be binary, we use unions and products in place of sequences, and we rewrite multiset and cycle using the pointing operator, Θ, where ΘA can be interpreted as pointing at any of the atoms of a structure A. Analytically, we have

$$C = \Theta A \qquad \Longrightarrow \qquad C(z) = \Theta A(z) \quad \text{where} \quad \Theta f(z) = z \cdot \frac{d}{dz} f(z).$$

Using the generating function from the above theorem, we see that if $A = \texttt{multiset}(B)$, then

$$\Theta A(z) = A(z) \times (\Theta B(z) + \Theta B(z^2) + \Theta B(z^3) + \cdots)$$

We will rewrite this expression using a new operator in the following way: ΘA marks an atom of an object of A. When $A = \texttt{multiset}(B)$, an object of A is a collection, with repetitions, of objects of B. Thus we can think of marking an atom of an A object as really marking the corresponding B object, so we would like to say that $\Theta A = A \times \Theta B$ (as in the labelled case). However, in the unlabelled case this is not quite accurate. We cannot distinguish between $\{b, b, c\}$ with the first b marked, and the same set with the second b marked at the same atom. So instead, we will think of marking not one b, but all of them "stacked" together. Thus, $\Theta A = \mathit{stack}(\Theta(B)) \times \texttt{multiset}(B)$.

We denote this "stacking" function, or the *diagonal*, by Δ. Then is is easily demonstrated that the generating function corresponding to Δ is given by the following.

THEOREM 2. $\Delta F(z) = F(z) + F(z^2) + F(z^3) + \cdots$

Then we can rewrite our expression for multiset as $A = \texttt{multiset}(B) \Rightarrow \Theta A = \Delta\Theta B \times A$, which is exactly what we started with.

Given a numeric sequence $\{u(k)\}_{k=1}^{\infty}$, the *generalised diagonal* is defined by

$$\Delta_{\{u(k)\}} = \sum_{k=1}^{\infty} u(k) \Delta^{(k)}.$$

Combinatorially, this means taking a weighted combination of diagonals. Analytically, this also defines a linear operator over generating functions involving a weighted combination of $A(z), A(z^2), \ldots$

$$C = \Delta_{\{u(k)\}} A \qquad \Longrightarrow \qquad C(z) = \Delta_{\{u(k)\}} A(z) = \sum_{k=1}^{\infty} u(k) A(z^k).$$

Thus, we can rewrite our expression for multiset as

$$A = \texttt{multiset}(B) \Rightarrow \Theta A(z) = A(z) \cdot \Delta_{\{1\}} \Theta B(z).$$

We also find that

$$A = \texttt{cycle}(B) \Rightarrow \Theta A(z) = \Delta_{\{\varphi(k)\}} \frac{\Theta B(z)}{1 - B(z)}.$$

For both multiset and cycle, we can derive similar relations when we have an imposed restriction on the cardinality, thus all cardinality restrictions are also expressible in terms of the union, product, generalised diagonal and pointing operations.

4. Generation

Once we know how to count the number of objects C_n in the decomposable class $\mathcal{C}$, we can generate objects of size n. The generation of most structures is the same here as in the labelled case, and so we will only briefly discuss the two main algorithms. Again, details can be found in [1].

Say $\mathcal{C} = \mathcal{A} \times \mathcal{B}$. Then $C_n = \sum_n A_k B_{n-k}$ and thus $\Pr(K = k) = a_k b_{n-k}/c_n$. To generate an object in $\mathcal{C}$ of size n using the *sequential* algorithm, we randomly pick a number between 0 and c_n, and then we sequentially, "from the left", add up the probabilities for increasing values of k until we find the interval in which our random number lies. For that k, we then recursively generate an element of size k in $\mathcal{A}$ and an object of size $n - k$ in $\mathcal{B}$.

Random generation of a product using the sequential algorithm corresponds to calculating the path length of the corresponding parse tree for the product, and thus is known to have worst case complexity $\mathcal{O}(n^2)$.

If instead, we use the *Boustrophedonic* algorithm, we can reduce the worst case complexity from $\mathcal{O}(n^2)$ to $\mathcal{O}(n \log n)$. For this algorithm, instead of starting "on the left" and sequentially adding up probabilities to determine which value of k to use, we first start at the left, then the right, and keep alternating back and forth until we have either found the value of k or the value of $n - k$, whichever is smaller. (*Boustrophedonic* means "turning like oxen in ploughing" (Webster).) If either value is small, we will find it quickly, and in the worst case (when $n = k$), the problem will be split into two subproblems of equal size, which is also an advantage. Analysis of this problem gives us the following theorem.

THEOREM 3 (BOUSTROPHEDONIC GENERATION, UNLABELLED CASE). *Any class of unlabelled structures that admits a (possibly recursive) specification in terms of the given constructions is susceptible to a random generation algorithm of arithmetic complexity* $\mathcal{O}(n \log n)$.

A cost analysis of both algorithms on average case problems suggests that while the Boustrophedonic algorithm has better worst case behaviour, in the sequential algorithm we can take advantage of optimisation strategies suggested by the cost calculus, and it appears that most classical classes of structures can be generated in time asymptotic to $\frac{1}{2} n \log n$ or sometimes even $\mathcal{O}(n)$ on average.

For unlabelled objects, we have the new problem of generating elements that use the Δ operator. Say $\mathcal{G} = \Delta\mathcal{F}$. Then $G_n = \sum_{k|n} F_{n/k}$, and so the probability that $\mathcal{G}$ produces a k-stack of size n is $F_{n/k}/G_n$. Now that we know the distribution of probabilities, we can use a sequential algorithm to choose the appropriate value for the stack size k, and hence generate the element. It is important to

note that generating $\Delta\mathcal{F}$ has the same complexity as generating $\mathcal{F}$, since we only need to consider at most the number of divisors of n.

5. Implementation

This system has been implemented in the Maple language by P. Zimmermann and E. Murray (for a detailed description of a preliminary version, see [2]). The program takes a grammar of an unlabelled (or labelled) structure and rewrites it into the standard, Chomsky-Normal form using only union, product, pointing and stacking. Then, for each non-terminal in the new grammar, it creates a function based on a pre-existing template to count objects of size n defined by the non-terminal, and another function to draw an object of size n defined by the non-terminal. These functions are stored in tables, and called when the user asks to count or generate an object of the original specification.

For example, with the grammar C = Union(A,B), the program will create a function gC to generate C objects:

gC := procedure(n : integer)
  U :=Uniform([0, 1]);
  if $U < a_n/c_n$ then gA(n) else gB(n) fi
end.

Some examples that have been implemented using this program include 2-3 trees, binary trees of fixed or bounded height, arithmetic expressions of one variable and circuits with resistors in parallel and series. For instance, the grammar specification for binary trees of height ≤ 3 is

$$\{B_0 = Z, B_1 = \mathrm{Union}(Z, \mathrm{Prod}(B_0, B_0)),$$
$$B_2 = \mathrm{Union}(Z, \mathrm{Prod}(B_1, B_1)), B_3 = \mathrm{Union}(Z, \mathrm{Prod}(B_2, B_2))\}.$$

6. Conclusion

This systematic approach to random generation not only handles widely different problems that have been studied in detail elsewhere on an individual basis, but it also in some cases improves the worst case bounds previously known. For example, the boustrophedonic algorithm gives $\mathcal{O}(n \log n)$ worst case time to all unambiguous context-free languages, whereas the best previous bound (due to Hickey and Cohen) is $\mathcal{O}(n^{2+\epsilon})$. Further areas of research involve extending the system to include powersets, and to consider the *unranking* problem: given a structure A, two integers n and $1 \leq i \leq A_n$, output the ith object of size n in A. If we could do this, we would be able to generate all objects of a given size very efficiently.

Bibliography

[1] Flajolet (Philippe), Zimmerman (Paul), and Van Cutsem (Bernard). – A calculus for the random generation of labelled combinatorial structures. *Theoretical Computer Science*, vol. 132, n° 1-2, 1994, pp. 1–35.

[2] Zimmermann (Paul). – Gaïa: A package for the random generation of combinatorial structures. *The Maple Technical Newsletter*, vol. 1, n° 1, Spring 1994.

Introduction to q-calculus

Laurent Habsieger
Université Bordeaux I

January 24, 1994

[summary by Xavier Gourdon]

Abstract

Many mathematical formulæ can be generalised by adding a new parameter q, leading to what is called a q-analogue, because the original formula can be obtained as the limit when q tends to 1. We present here a combinatorial introduction to the q-calculus.

1. Partitions and words

1.1. Partitions. A partition λ is a decreasing sequence $(\lambda_1, \ldots, \lambda_k)$ of positive integers: $\lambda_i \in \mathbb{N}^*$ and $\lambda_i \geq \lambda_{i+1}$ for all i. The length of λ is $\ell(\lambda) = k$, its height is $|\lambda| = \sum_{i=1}^{k} \lambda_i$. If $|\lambda| = n$, we say that λ is a partition of n. For all $\ell, m \in \mathbb{N}$, let

$$P(\ell, m) = \{\lambda : \ \ell(\lambda) \leq \ell \quad \text{and} \quad \lambda_1 \leq m\}.$$

It is possible to determine a partition λ from the numbers $m_i = \mathrm{Card}\{j : \ \lambda_j = i\}$ denoting the multiplicity of i in λ. In this way, the partition λ can be written as $\lambda = (1^{m_1} 2^{m_2} \cdots)$.

A nice way to represent a partition λ is to use its Ferrers diagram

$$D_\lambda = \{(i, j) \in \mathbb{Z}^2 : \ 1 \leq i \leq \lambda_j \quad \text{and} \quad 1 \leq j \leq \ell(\lambda)\}.$$

The number $P(\ell, m)$ can be viewed as $P(\ell, m) = \{\lambda : \ D_\lambda \subset (m^\ell)\}$. The conjugate partition of λ is the partition λ' whose Ferrers diagram is symmetric from D_λ with respect to the first bisecting line. We have $|\lambda'| = |\lambda|$ and $(\lambda')' = \lambda$.

1.2. Gaussian polynomials. Ferrers diagrams enable to establish a correspondence between partitions of $P(\ell, m)$ and paths joining $(0, \ell)$ to $(m, 0)$ with the steps $(0, -1)$ and $(1, 0)$. Such paths have $m + \ell$ steps (m horizontal and ℓ vertical) so $\mathrm{Card}\, P(\ell, m) = \binom{m+\ell}{m}$. To take into account the height in this statistic, we introduce its generating function with respect to a new variable q. We have

$$\sum_{\lambda \in P(\ell, m)} q^{|\lambda|} = \frac{(q)_{m+\ell}}{(q)_m (q)_\ell} \quad \text{where} \quad \begin{cases} (q)_k = \prod_{i=1}^{k} (1 - q^i) & k \geq 1, \\ (q)_0 = 1. \end{cases} \tag{1}$$

Letting $q \to 1$ in this identity, we find again $\mathrm{Card}\, P(\ell, m) = \binom{m+\ell}{m}$. This motivates the definition of a q-analogue of the binomial coefficients, denoted by

$$\begin{bmatrix} m + \ell \\ \ell \end{bmatrix} = \frac{(q)_{m+\ell}}{(q)_m (q)_\ell},$$

and called Gaussian polynomials. They satisfy several q-properties like Pascal recurrences or symmetry.

By letting $\ell \to \infty$ in identity (1), we get

$$\sum_{\lambda:\ \lambda_1 \le m} q^{|\lambda|} = \frac{1}{(q)_m} = \sum_{\lambda:\ \ell(\lambda) \le m} q^{|\lambda|}. \tag{2}$$

This last equality is obtained from the conjugate partitions. Then letting $m \to \infty$, we find

$$\sum_{\lambda} q^{|\lambda|} = \sum_{n \ge 0} p(n) q^n = \frac{1}{(q)_\infty}, \tag{3}$$

where $p(n)$ is the total number of partitions of n.

1.3. Infinite products. Formula (2) can be refined by introducing a new variable x. More precisely, denoting

$$(x)_\infty = \prod_{i=0}^{\infty} (1 - xq^i),$$

we have the identity

$$\frac{1}{(x)_\infty} = \sum_{m_0, m_1, \ldots} \left(\prod_{i=0}^{\infty} x^{m_i} q^{i m_i} \right) = \sum_{\ell \ge 0} x^\ell \sum_{\lambda:\ \ell(\lambda) \le \ell} q^{|\lambda|} = \sum_{\ell \ge 0} \frac{x^\ell}{(q)_\ell}.$$

In the same vein, by expanding $(-x)_\infty$ we have

$$(-x)_\infty = \sum_{\ell \ge 0} \frac{q^{\binom{\ell}{2}}}{(q)_\ell} x^\ell.$$

These two identities are sometimes called Euler identities.

Jacobi identity. The triple product identity of Jacobi is

$$(q)_\infty (x)_\infty (qx^{-1})_\infty = \sum_{n \in \mathbb{Z}} (-1)^n x^n q^{\binom{n}{2}}.$$

As a corollary, we have the formulæ

$$(q)_\infty = \sum_{n \in \mathbb{Z}} (-1)^n q^{n(3n+1)/2} \tag{4}$$

$$(q)_\infty^3 = \sum_{n \in \mathbb{N}} (-1)^n (2n+1) q^{n(n+1)/2}. \tag{5}$$

The first one is Euler's pentagonal number theorem, and can be used with (3) to establish several congruences relations satisfied by the partition numbers $p(n)$.

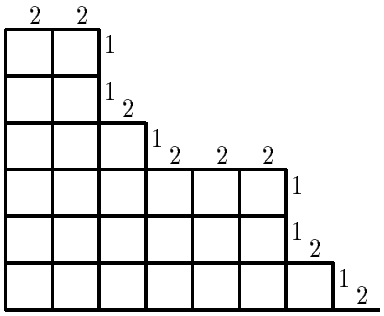

FIGURE 1. With $\ell = 6$ and $m = 9$, the partition $\lambda = (8, 7, 7, 4, 2, 2)$ is associated to the word $w = 221122122211212$.

2. Words

2.1. Correspondence between partitions and binary words. Consider $\lambda \in P(\ell, m)$. Its Ferrers diagram can be considered as a path joining points $(0, \ell)$ and $(m, 0)$ by m horizontal steps and ℓ vertical steps. We encode this path with a word on $\{1, 2\}^*$, associating a 1 for each vertical step, a 2 for each horizontal step (see figure 1). This construction defines a correspondence between $P(\ell, m)$ and $M(\ell, m)$, the words of $\{1, 2\}^*$ with ℓ "1" and m "2".

We define the number of inversions of a word w in $\{1, 2\}^*$ by

$$\text{Inv}\, w = \text{Card}\{(i, j) : \ 1 \le i < j \le \ell + m \quad \text{and} \quad 2 = w_i > w_j = 1\}. \tag{6}$$

We have $\text{Inv}\, w = |\lambda|$, where w is the word obtained from λ by the correspondence, thus

$$\sum_{w \in M(\ell,m)} q^{\text{Inv}(w)} = \begin{bmatrix} m + \ell \\ \ell \end{bmatrix}.$$

Another interesting parameter is the major index defined by

$$\text{Maj}\, w = \sum_{w_i > w_{i+1}} i, \tag{7}$$

and surprisingly, its generating function is the same as the one of Inv.

2.2. Statistics on words over n letters. The previous discussion finds a natural generalisation by considering $M(a_1, \ldots, a_n)$, the set of words with n letters where the i-th letter appears exactly a_i times. The length of such a word w is $a_1 + \cdots + a_n$. The parameters $\text{Inv}\, w$ and $\text{Maj}\, w$ are defined as in (6) and (7). The Z-statistic (called like this because of Zeilberger work [2]) of a word w is defined as

$$z(w) = \sum_{1 \le i < j \le n} \text{Maj}\, w_{i,j}$$

where $w_{i,j}$ is the word obtained from w by keeping only the i-th and the j-th letter. These parameters satisfy

$$\sum_{w \in M(a_1,\ldots,a_n)} q^{\text{Inv}(w)} = \sum_{w \in M(a_1,\ldots,a_n)} q^{\text{Maj}(w)} = \sum_{w \in M(a_1,\ldots,a_n)} q^{z(w)} = \begin{bmatrix} a_1 + \cdots + a_n \\ a_1, \ldots, a_n \end{bmatrix} := \frac{(q)_{a_1 + \cdots + a_n}}{(q)_{a_1} \cdots (q)_{a_n}},$$

providing a q-analogue of multinomial coefficients.

3. Basic hypergeometric functions

We use the notations

$$(a)_\infty = (a;q)_\infty = \prod_{i=0}^{\infty}(1-aq^i), \qquad (a)_n = (a;q)_n = \frac{(a)_\infty}{(aq^n)_\infty}$$

and we define the basic hypergeometric series as

$${}_r\phi_s\left(\begin{array}{c}\alpha_1,\ldots,\alpha_r\\ \beta_1,\ldots,\beta_s\end{array};x\right) = \sum_{n=0}^{\infty}\frac{(\alpha_1)_n\cdots(\alpha_r)_n}{(q)_n(\beta_1)_n\cdots(\beta_r)_n}x^n.$$

The $q\to 1$ limit in this expression leads to a classical hypergeometric series, thus we have defined a q-analogue of hypergeometric series. A good survey of basic hypergeometric series can be found in [3].

3.1. The q-binomial theorem. The relation $(1-x){}_1\phi_0(a;x) = (1-ax){}_1\phi_0(a;qx)$ together with ${}_1\phi_0(a;0)=1$ leads to the q-binomial theorem:

$$ {}_1\phi_0(a;x) = \frac{(ax)_\infty}{(x)_\infty}. \tag{8}$$

By setting $a=q^{-n}$ then $x\to -xq^{-n}$, we deduce

$$\prod_{i=0}^{n-1}(1+q^ix) = \sum_{k=0}^{n}\begin{bmatrix}n\\k\end{bmatrix}q^{\binom{k}{2}}x^k.$$

When $q\to 1$, this leads to the classical binomial theorem.

3.2. Heine transforms. Like classical hypergeometric functions, the basic hypergeometric functions satisfy several identities. A first family is the Heine transforms:

$${}_2\phi_1\left(\begin{array}{c}\alpha,\beta\\ \gamma\end{array};x\right) = \frac{(\beta)_\infty(\alpha x)_\infty}{(\gamma)_\infty(x)_\infty}\,{}_2\phi_1\left(\begin{array}{c}\gamma/\beta,x\\ \alpha x\end{array};\beta\right)$$

$${}_2\phi_1\left(\begin{array}{c}\alpha,\beta\\ \gamma\end{array};x\right) = \frac{\left(\frac{\alpha\beta}{\gamma}x\right)_\infty}{(x)_\infty}\,{}_2\phi_1\left(\begin{array}{c}\gamma/\alpha,\gamma/\beta\\ \gamma\end{array};\frac{\alpha\beta}{\gamma}x\right).$$

3.3. Pfaff-Saalschütz q-theorem. This result applies to functions of the type ${}_3\phi_2$. For all non-negative integer n, we have

$${}_3\phi_2\left(\begin{array}{c}a,b,q^{-n}\\ c,\frac{ab}{c}q^{1-n}\end{array};q\right) = \frac{(c/a)_n(c/b)_n}{(c)_n(c/ab)_n}. \tag{9}$$

There exists several equivalent forms of this theorem. By letting $n\to+\infty$ in (9), we obtain the Gauss q-theorem

$${}_2\phi_1\left(\begin{array}{c}a,b\\ c\end{array};\frac{c}{ab}\right) = \frac{(c/a)_\infty(c/b)_\infty}{(c)_\infty(c/ab)_\infty}.$$

Another corollary of the Pfaff-Saalschütz q-theorem is the q-formula of Chu-Vandermonde, obtained by setting $a=q^{n+1}$, $b=q^{-k}$ and $c=q^{m+1}$ in (9)

$$\sum_{i=0}^{k}(-1)^iq^{i(i-1)/2+(m-n)i}\begin{bmatrix}n+i\\i\end{bmatrix}\begin{bmatrix}m+k\\k-i\end{bmatrix} = \begin{bmatrix}k+m-n-1\\k\end{bmatrix}.$$

There exists several generalizations of the Pfaff-Saalschütz q-theorem. One is called the Dougall q-theorem, it applies to functions of the type ${}_8\phi_7$.

4. q-analogues of usual tools

4.1. q-derivative. The q-derivative of a function f is defined as

$$\delta_q f(t) = \frac{f(t) - f(qt)}{(1-q)t}.$$

The formulæ of the classical derivative have their q-analogues with respect to the q-derivative.

4.2. q-integration. The function $g(t) = \int_0^t f(x)\, d_q x$ must satisfy $\delta_q g = f$, so

$$\begin{aligned} g(t) - g(qt) &= t(1-q)f(t) \\ g(qt) - g(q^2t) &= qt(1-q)f(qt) \\ \cdots &= \cdots \end{aligned}$$

thus $g(t) = g(t) - g(0) = \sum_{n\geq 0} q^n t(1-q) f(q^n t)$, and we define

$$\int_0^t f(x)\, d_q x = t(1-q)\sum_{n\geq 0} q^n f(q^n t).$$

Like the classical integral, there exists a q-formula of integration by parts. There are several ways of defining an improper integral, for example

$$\int_0^{+\infty} f(t)\, d_q t = (1-q)\sum_{n\in\mathbb{Z}} q^n f(q^n) \quad \text{and} \quad \int_0^{+\infty} f(t)\, d_q t = \int_0^{1/(1-q)} f(t)\, d_q t$$

are q-analogues of $\int_0^{+\infty} f(t)\, dt$.

4.3. q-differential equations. The q-differential equation $\delta_q f(t) = f(t)$ admits the solution

$$f(t) = \frac{f(qt)}{1-t(1-q)} = \cdots = \frac{f(0)}{(t(1-q))_\infty},$$

thus the solution f with $f(0) = 1$ is

$$e_q(t) = \frac{1}{(t(1-q))_\infty} = \sum_{n\geq 0} \frac{(1-q)^n}{(q)_n t^n},$$

(the last identity is obtained from the q-binomial theorem (8) with $a = 0$ and $x = t(1-q)$) providing a q-analogue of the expansion of $\exp(t)$.

As for the q-differential equation $\delta_q f(t) = f(qt)$, the solution which takes the value 1 at 0 is

$$E_q(t) = (-t(1-q))_\infty = \sum_{n\geq 0} q^{\binom{n}{2}} \frac{(1-q)^n}{(q)_n} t^n,$$

the last identity being a consequence of the q-binomial theorem applied with $a = -t(1-q)/x$ and $x \to 0$. This second q-analogue of the expansion of $\exp(t)$ satisfy the obvious relation $e_q(t)E_q(-t) = 1$. Nevertheless, there does not exist any simple relation between $e_q(x)e_q(y)$, $E_q(x)E_q(y)$ and

$e_q(x+y)$, $E_q(x+y)$. A q-analogue of the relation $\exp(x+y) = \exp(x)\exp(y)$ is given by the formula

$$e_q(x)E_q(y) = \sum_{n=0}^{+\infty} \frac{\prod_{k=0}^{n-1}(x+q^k y)}{\prod_{k=1}^{n}\frac{1-q^k}{1-q}},$$

obtained from the q-binomial theorem with $a = -y/x$ and $x = x(1-q)$.

4.4. The q-gamma function. The q-gamma function is defined as

$$\Gamma_q(s) = \frac{(q)_\infty}{(q^s)_\infty}(1-q)^{1-s}, \qquad s \in \mathbb{C} \setminus \{0, -1, -2, \dots\},$$

which tends to $\Gamma(s)$ as $q \to 1$. The functional equation of Γ_q is

$$\Gamma_q(s+1) = \frac{1-q^s}{1-q}\Gamma_q(s),$$

and since $\Gamma_q(1) = 1$, we have for all positive integer n

$$\Gamma_q(n+1) = \prod_{k=1}^{n}\frac{1-q^k}{1-q} = \frac{(q)_n}{(1-q)^n},$$

which is a q-analogue of $\Gamma(n+1) = n!$. The function $\log\Gamma_q(x)$ is convex for $x > 0$. An integral representation of Γ_q is

$$\Gamma_q(s) = \int_0^{1/(1-q)} t^{s-1}E_q(-qt)\,d_q t.$$

There also exists a q-analogue of the Gauss duplication formula.

4.5. The q-beta function. An equivalent form of the q-binomial theorem is

$$\int_0^1 t^{x-1}\frac{(qt)_\infty}{(q^y t)_\infty}\,d_q t = \frac{\Gamma_q(x)\Gamma_q(y)}{\Gamma_q(x+y)}, \qquad (\Re(x) > 0).$$

This expression can be used to define the most well known and the most useful q-analogue of the beta function.

Bibliography

[1] Andrews (George E.). – *The Theory of Partitions.* – Addison-Wesley, 1976, *Encyclopedia of Mathematics and its Applications*, vol. 2.

[2] Bressoud (D. M.) and Zeilberger (D.). – A proof of Andrew's q-Dyson conjecture. *Discrete Mathematics*, vol. 54, 1985, pp. 201–224.

[3] Gasper (George) and Rahman (Mizan). – *Basic Hypergeometric Series.* – Cambridge University Press, 1990, *Encyclopedia of Mathematics and its Applications*, vol. 35.

Overlap-Free Words

Julien Cassaigne
ENS Paris

November 29, 1993

[summary by Philippe Dumas]

An overlap-free word is a word without overlapping of two distinct occurrences of a factor. As an example the word *abaab* is overlap-free, but the word $ab\underline{bbab}\overline{bbab}aab$ is not because the factor *bbab* occurs twice and the occurrences overlap by *b*. The first result about these words is due to Axel Thue [10]. Indeed the infinite word now referred to as the Thue-Morse word has no overlapping factor [8, p. 23]. Recall that this word may be defined as a fixed point $t = \vartheta^\omega(a)$ of the substitution ϑ which maps *a* and *b* onto *ab* and *ba* respectively. The proof depends upon the fact that $\vartheta(w)$ has no overlapping factor if the word w does not.

We want to study the number u_n of overlap-free words of length n over a two letter alphabet $\{a, b\}$. All the factors of the Thue-Morse word t are overlap-free words. This yields the inequality $u_n \geq t_n$, where t_n is the number of factors of t with length n. This number t_n is known to be [2, 5]

$$t_n = \begin{cases} 4n - 2.2^k & \text{if} \quad 2.2^k \leq n \leq 3.2^k, \\ 2n + 4.2^k & \text{if} \quad 3.2^k \leq n \leq 4.2^k. \end{cases}$$

On the other hand, Restivo and Salami [9] have shown the asymptotic inequality[1]

$$u_n \preccurlyeq n^{\log_2 15}.$$

Kobayashi [7] improved these results by the estimate

$$n^{1.155} \preccurlyeq u_n \preccurlyeq n^{1.587}.$$

Overlap-free words which can be extended to infinity in both directions are factors of the Thue-Morse word. Counting the words extensible to the right gives the lower bound. Thanks to ϑ, one can build overlap-free words of length $2n$ from overlap-free words of length n, hence the upper bound. Carpi [3] gave a way to obtain upper bounds n^r arbitrarily near the optimal value, but gave no numerical result. In addition he pointed out that one can compute the sequence u_n by a finite automaton.

1. Linear representation

Decomposition. All the cited results rely upon a decomposition of the overlap-free words. Apart from a set S of seventy-six little words, every word which is overlap-free and of length greater than 3, or which has a unique overlapping by only one letter (J. Cassaigne uses the term of minimal overlap for these words $xuxux$, where x is a letter and u a word), can be written in a unique manner in the form $r_1\vartheta(u)r_2$ with u an overlap-free word and r_1, r_2 are in $\{\varepsilon, a, b, aa, bb\}$. J. Cassaigne

[1]The symbol $\preccurlyeq$ has the meaning given by Bourbaki. In Landau notation, it is a big O.

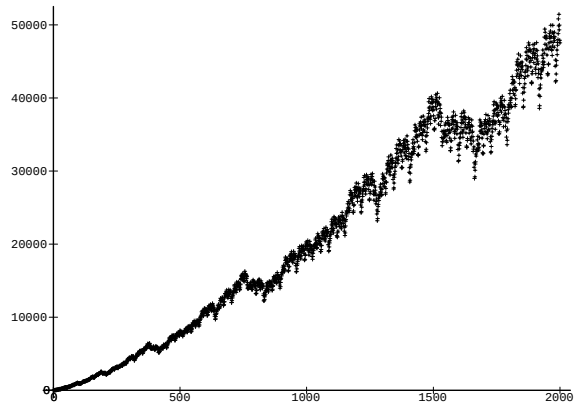

FIGURE 1. The number of overlap-free words presents a behaviour of polynomial type, but with heavy oscillations.

prefers to use a subtler decomposition, the advantage of which is to make independent both ends of the word.

He defines an action of the monoid $G = (E \times E)^*$, with $E = \{\delta, \iota, \kappa\}$, on the set $\{a, b\}^*$. The key idea is to modify the word $\vartheta(u)$ by deleting or inverting (in the sense that the inverse of a is b and conversely) letters at the ends of the word, in such a way that the resulting word is almost overlap-free if w is overlap-free and that every overlap-free word can be reached from S by this action. To be more precise, let U be the set of overlap-free words of length greater than 3, V be the set of words $xuxux$ with only an overlapping of only one letter x. J. Cassaigne proved the following result.

LEMMA 1. *Every word $w \in (U \cup V) \setminus S$ can be written $w = v.\gamma$ in a unique manner with $\gamma \in E \times E$ and $v \in U \cup V$. Moreover the length of v is less than the length of w.*

Here is why V is adequate. As a matter of fact, the word w may be overlap-free but v may be a minimal overlap.

Next, it is possible to iterate the process. Starting from an element $s \in S$, one applies a element $g \in G$ to obtain $z = s.g$. Technically, one has to remove some little words, so one uses only a subset Z of $S \times G$.

LEMMA 2. *Every word $w \in U \cup V$ has a unique decomposition $w = s.g$ with $(s, g) \in Z$.*

Automaton. The next crucial step is to consider the languages L and M of the words $(s, g) \in Z$ such that $s.g$ lies in U or V.

PROPOSITION 1. *The languages L and M are rational.*

Moreover, these languages are recognized by the same automaton, with only a change of the set of terminal states. The proof relies on the study of the overlappings of $s.g.\gamma$ with respect to γ and the overlappings of $s.g$. The states of the automaton are 2-tuples (i, j) with $i, j \in \{1, \ldots, 4\}$, which translate the prefix and suffix forms of the words. For example a word is of type $(2, 3)$ if it starts with *abb* or *baa* and ends with *abaa* or *babb*. The transitions of the automaton are labelled by the

elements of $E \times E$ and are expressed by $(i,j).(\gamma_1,\gamma_2) = (\varphi(i,\gamma_1),\varphi(j,\gamma_2))$. As we predicted the prefix and suffix of the words are disconnected. Roughly speaking, the automaton is made from three blocks: an initial block, which depends on the form of the words from S, and two blocks of the same structure, one which accepts the words in L and the other which accepts the words in M.

The automaton may be viewed in terms of matrices. Let $U_n(i,j)$, $V_n(i,j)$ be the number of words of length n in U and V respectively. Let δ, ι, κ be the matrix of the transformations δ, ι, κ,

$$\delta = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \iota = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \kappa = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

PROPOSITION 2. *The matrices U_n and V_n satisfy the following recurrence,*

$$\begin{cases} V_{2n} &= 0, \\ V_{2n+1} &= \kappa {V_{n+1}}^t \delta + \delta {V_{n+1}}^t \kappa, \\ U_{2n} &= \iota {V_n}^t \iota + \delta {V_{n+1}}^t \delta + (\kappa+\iota) {U_n}^t (\kappa+\iota) + \delta {U_{n+1}}^t \delta, \\ U_{2n+1} &= \iota {V_{n+1}}^t \delta + \delta {V_{n+1}}^t \iota + (\kappa+\iota) U_{n+1}^{t\delta} + \delta {U_{n+1}}^t (\kappa+\iota), \end{cases}$$

for any integer n greater than 3.

The first few values,

$$V_4 = 0, \quad V_5 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad U_4 = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad U_5 = \begin{pmatrix} 0 & 2 & 2 & 0 \\ 2 & 0 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix},$$

enables computation of the matrices U_n and V_n, and hence

$$u_n = \sum_{1 \le i,j \le 4} U_n(i,j).$$

As a consequence of this recurrence, we have the following assertion.

THEOREM 1. *The sequence (u_n) of the number of overlap-free words is 2-regular.*

Recall the notion of B-regularity due to Allouche and Shallit [1].

2. Asymptotic behaviour

As a 2-regular sequence, u_n grows polynomially and it is natural to search for the best numbers α and β such that

$$n^{\alpha-\varepsilon} \preccurlyeq u_n \preccurlyeq n^{\beta+\varepsilon},$$

for any positive ε. Using integers the binary expansion of which has a given pattern (namely the numbers 7.2^k and $(22.4^k - 1)/3$), J. Cassaigne gives some possible values for α and β but not the best ones. He obtains, with the previous results of Kobayashi, the following inequalities for the best α and β

$$1.155 < \alpha < 1.276 < 1.332 < \beta < 1.587.$$

The sums

$$s_n = \sum_{k=0}^{n-1} u_k$$

are simpler to study because they give an increasing sequence; consideration of the integers $n = 7.2^k$ suffices to obtain the behaviour of s_n.

THEOREM 2. *The sequence s_n grows like $n^{\log_2 \zeta}$, where*

$$\zeta = \sqrt{3} + \frac{3 + \sqrt{5 + 4\sqrt{3}}}{2}.$$

As a matter of fact, the number ζ is the largest eigenvalue of a linear representation of the 2-regular sequence s_n. The result is not surprising because this is the situation in all the known cases [6], although there is no proof in the general case.

Bibliography

[1] Allouche (J.-P.) and Shallit (J.). – The ring of k-regular sequences. *Theoretical Computer Science*, vol. 98, 1992, pp. 163–197.

[2] Brlek (S.). – Enumeration of factors in the Thue-Morse word. *Discrete Applied Mathematics*, 1989, pp. 83–96.

[3] Carpi (Arturo). – Overlap-free words and finite automata. *Theoretical Computer Science*, vol. 115, 1993, pp. 243–260.

[4] Cassaigne (Julien). – Counting overlap-free binary words. In Enjalbert (Patrice), Finkel (A.), and Wagner (Klaus W.) (editors), *STACS '93. Lecture Notes in Computer Science*, vol. 665, pp. 216–225. – Würzburg, 1993.

[5] de Luca (A.) and Varricchio (S.). – Some combinatorial properties of the Thue-Morse sequence and a problem in semigroups. *Theoretical Computer Science*, vol. 63, 1989, pp. 333–348.

[6] Flajolet (Philippe), Grabner (Peter), Kirschenhofer (Peter), Prodinger (Helmut), and Tichy (Robert). – Mellin transforms and asymptotics: Digital sums. *Theoretical Computer Science*, vol. 123, n° 2, 1994, pp. 291–314.

[7] Kobayashi (Y.). – Enumeration of irreducible binary words. *Discrete Applied Mathematics*, vol. 20, 1988, pp. 221–232.

[8] Lothaire (M.). – *Combinatorics on Words.* – Addision–Wesley, 1983, *Encyclopedia of Mathematics and its Applications*, vol. 17.

[9] Restivo (A.) and Salemi (S.), Nivat (M.) and Perrin (D.) (editors). – *Overlap-free words on to symbols*, pp. 198–206. – Springer–Verlag, 1985, *Lectures Notes in Computer Science*, vol. 192.

[10] Thue (A.). – Über unendliche Zeichenreihen. *Norske Vid. Selsk. Skr. I. Mat. Nat. Kl.*, n° 7, 1906, pp. 1–67. – Also in [11].

[11] Thue (A.). – *Selected Mathematical Papers.* – Universitetforlaget, Oslo, 1977. Edited by T. Nagell, A. Selberg, S. Selberg and K. Thalberg.

Descents in Words

Jean-Marc Fédou
LaBRI, Université de Bordeaux I

March 28, 1994

[summary by Dominique Gouyou-Beauchamps]

1. Introduction

Let S_n denote the symmetric group on $\{1, 2, \ldots, n\}$. For a permutation $\sigma \in S_n$, the *rise set*, *descent set*, *inversion set*, and their cardinalities are respectively defined by

$$\begin{aligned} \operatorname{Ris}\sigma &= \{i : 1 \le i \le n-1, \sigma(i) < \sigma(i+1)\}, & \operatorname{ris}\sigma &= |\operatorname{Ris}\sigma|, \\ \operatorname{Des}\sigma &= \{i : 1 \le i \le n-1, \sigma(i) > \sigma(i+1)\}, & \operatorname{des}\sigma &= |\operatorname{Des}\sigma|, \\ \operatorname{Inv}\sigma &= \{(k,m) : 1 \le k < m \le n, \sigma(m) < \sigma(k)\}, & \operatorname{inv}\sigma &= |\operatorname{Inv}\sigma|. \end{aligned}$$

The *set of common descents* of a pair of permutations $(\sigma_1, \sigma_2) \in S_n^2$ is defined as $\operatorname{DD}(\sigma_1, \sigma_2) = \operatorname{Des}\sigma_1 \cap \operatorname{Des}\sigma_2$ and one notes $\operatorname{dd}(\sigma_1, \sigma_2) = |\operatorname{DD}(\sigma_1, \sigma_2)|$.

Now we recall three definitions of Bessel functions (with $(q)_n = (1-q)(1-q^2)\cdots(1-q^n)$):

$$J_\nu(x) = \sum_{n\ge 0} \frac{(-1)^n \left(\frac{1}{2}x\right)^{2n+\nu}}{n!\Gamma(\nu+n+1)}, \qquad \text{(usual)}$$

$$J_\nu(x) = \sum_{n\ge 0} \frac{(-1)^n x^{n+\nu}}{n!(n+\nu)!}, \qquad \text{(combinatorial)}$$

$${}^qJ_\nu(x) = \sum_{n\ge 0} \frac{(-1)^n q^{\binom{n+\nu}{2}} x^{n+\nu}}{(q)_n (q)_{n+\nu}}, \qquad (q\text{-analog}).$$

Perhaps the first combinatorial context for an element of $\{J_\nu\}_{\nu\ge 0}$ was discovered by L. Carlitz, R. Scoville, and T. Vaughan [4].

THEOREM 1 ([4, 12]). *The generating function of Bessel type for the sequence of polynomials*

$$a_n(y,q) = \sum_{(\alpha,\beta)\in S_n^2} q^{\operatorname{inv}\alpha} q^{-\operatorname{inv}\beta} y^{\operatorname{dd}(\alpha,\beta)} \qquad \textit{is} \qquad \sum_{n\ge 0} q^{\binom{n}{2}} \frac{a_n(y,q)}{(q)_n (q)_n} x^n = \frac{1-y}{{}^qJ_0(x(1-y)) - y}.$$

Setting $q = 1$ and $y = 0$ in Theorem 1, they deduce that the coefficient $a_n(0,1)$ of $x^n/(n!n!)$ in the series expansion of $1/J_0(x)$ is equal to the number of permutation pairs with no common descents.

First we remark that

$$\sum_{n\ge 0} q^{\binom{n}{2}} \frac{a_n(y,q)}{(q)_n (q)_n} x^n = \frac{1}{1 + \sum_{n\ge 1} (-1)^n (1-y)^{n-1} q^{\binom{n}{2}} x^n/(q)_n^2}. \tag{1}$$

Second we remark that an inversion formula such as (1) involving alternate sums of $(1-y)^{n-1}x^n$ is also present in the well known q-Eulerian polynomials and q-Euler polynomials [2, 3, 16, 17]. More precisely, let $A_n(y,q)$ denote the q-Eulerian polynomials. Then

$$\sum_{n\geq 0} A_n(y,q)\frac{x^n}{(q)_n} = \frac{1}{1+\sum_{n\geq 1}(-1)^n(1-y)^{n-1}x^n/(q)_n}.$$

A natural problem is to give for generating functions of the type

$$F(x,y) = \frac{1}{1+\sum_{n\geq 1}(-1)^n\lambda_n(1-y)^{n-1}x^n}$$

a combinatorial interpretation similar to the one described for the generating function $F(x,0)$ of heaps of pieces [18]. There are a number of powerful theories of inversion [13, 14, 17, 19] for dealing with combinatorial objects having generating functions of type $F(x,0)$. Using two such inversion formulas, we present new derivations of R. P. Stanley's generating functions for generalized q-Eulerian and q-Euler polynomials on r-uples of permutations [17]. We further indicate how one of the inversion formulas gives V. Diekert's lifting to the free monoïd of an inversion theorem of P. Cartier and D. Foata [5, 7]. The inversion theorems we use enumerate words in the free monoïd by adjacencies.

2. From the free to the trace monoïd

Let X be an alphabet. The empty word will be denoted by 1. The set of all words formed with letters in X by means of the concatenation product is known as the *free monoïd* generated by X and is denoted by X^*. In $\mathbb{Z}\langle\langle X\rangle\rangle$, the ring of formal power series of words in X^* with integer coefficients, the following inversion formula holds: $X^* = 1/(1-X)$.

Let θ be an irreflexive symmetric binary relation on X. Define $\equiv_\theta$ to be the binary relation (induced by θ) on X^* consisting of the set of pairs (w,v) of words such that there is a sequence $w = w_0, w_1, \ldots, w_m = v$ where each w_i is obtained by transposing a pair of letters in w_{i-1} that are consecutive and contained in θ.

Clearly, $\equiv_\theta$ is an equivalence relation on X^*. The quotient of X^* by $\equiv_\theta$ gives the *partially commutative monoïd* (or *trace monoïd*) induced by θ and denoted by $M(X,\theta)$. The equivalence class $\hat{w}$ of $w \in X^*$ is referred to as the *trace* of w.

A word $w = x_1x_2\cdots x_n \in X^*$ is said to be a *basic monomial* if $x_i\theta x_j$ for all $i \neq j$. Note that all the letters of a basic monomial are distinct. A trace $\hat{w}$ is said to be θ-*trivial* if any one of its representatives is a basic monomial. Letting $\mathcal{T}^+(X,\theta)$ be the set of θ-trivial traces, the inversion formula of Möbius type reads as follows.

Theorem 2 (P. Cartier and D. Foata). *For θ an irreflexive symmetric binary relation on X, the traces in $M(X,\theta)$ are generated by*

$$\sum_{\hat{w}\in M(X,\theta)} \hat{w} = \frac{1}{1+\sum_{\hat{t}\in\mathcal{T}^+(X,\theta)}(-1)^{l(\hat{t})}\hat{t}},$$

where $l(\hat{t})$ denotes the length of any representative of $\hat{t}$.

In terms of heaps of pieces, the Cartier-Foata's theorem is nothing but the inversion lemma for heap monoïd [18, prop. 5.1].

A natural question to ask is whether $\hat{w}$ and $\hat{t}$ can be replaced by some canonical representatives so that Theorem 2 remains true as a formula in the free monoïd X^*. As resolved by V. Diekert [6, 7], such canonical representatives exist if and only if θ admits a transitive orientation.

To be precise, a subset $\vec{\theta}$ of θ is an *orientation* of θ if θ is a disjoint union of $\vec{\theta}$ and $\{(x,y) : (y,x) \in \vec{\theta}\}$. The set of $t = t_1 t_2 \cdots t_n \in X^*$ satisfying $t_1 \vec{\theta} t_2 \vec{\theta} \cdots \vec{\theta} t_n$ is denoted by $T^+(X, \vec{\theta})$. Note that $T^+(X, \vec{\theta})$ is a set of representatives for the θ-trivial traces $\mathcal{T}^+(X, \theta)$ whenever $\vec{\theta}$ is transitive. A word $w = x_1 x_2 \cdots x_n \in X^*$ is said to have a *$\vec{\theta}$-adjacency* in position k if $x_k \vec{\theta} x_{k+1}$. We denote the number of $\vec{\theta}$-adjacencies of w by $\vec{\theta}\operatorname{adj} w$. Although V. Diekert did not explicitly introduce the notion of a $\vec{\theta}$-adjacency, his lifting theorem may be paraphrased as follows.

THEOREM 3 (V. DIEKERT). *Let θ be an irreflexive symmetric binary relation on X and $\vec{\theta}$ be an orientation of θ. Then, $\vec{\theta}$ is transitive if and only if there exists a complete set W of representatives for the traces of $M(X, \theta)$ such that*

$$\sum_{w \in W} w = \frac{1}{1 + \sum_{t \in T^+(X, \vec{\theta})} (-1)^{l(t)} t}.$$

Moreover, $W = \{w \in X^* : \vec{\theta}\operatorname{adj} w = 0\}$.

3. Descents in a word

Now X is a totally ordered alphabet. We say that a word $w = x_1 \cdots x_i x_{i+1} \cdots x_n$ of X^* has a *θ-descent* in position i when $x_i \theta x_{i+1}$ and $x_i > x_{i+1}$. We note $x \gg_\theta y$ (resp. $x \ll_\theta y$) when $x \theta y$ and $x > y$ (resp. $x < y$). Let $I^+ = \{x_1 \cdots x_n \in X^*, n > 0, x_1 \gg_\theta x_2 \gg_\theta \cdots \gg_\theta x_{n-1} \gg_\theta x_n\}$. Let $w \in X^*$, we denote by $\theta\operatorname{des}(w)$ the number of its θ-descents.

THEOREM 4 ([9]). *The following equality holds in the free monoïd*

$$\sum_{w \in X^*} y^{\theta \operatorname{des}(w)} w = \frac{1}{1 - \sum_{t \in I^+} (y-1)^{|t|-1} t}. \tag{2}$$

When $\gg_\theta$ is transitive, setting $y = 0$ in (2) gives the lifting of Theorem 2 to the free monoïd as stated in Diekert's Theorem 3. We close this section with two examples.

EXAMPLE (TRANSITIVE CASE). Let $X = \{a, b, c\}$, $a < b < c$ with $\theta = \{(a,b), (b,a), (a,c), (c,a)\}$. The θ-descents of a word correspond to factors ba and ca. Then $\gg_\theta$ is a transitive relation. Note that $I^+ = \{a, b, c, ba, ca\}$ is a complete set of the representatives for the θ-trivial traces $\mathcal{T}^+(X, \theta)$. From (2), we have

$$\sum_{w \in X^*} y^{\theta \operatorname{des}(w)} w = \frac{1}{1 - (a+b+c) + (1-y)(ba+ca)}.$$

Setting $y = 0$ gives an identity that can be viewed as having been lifted from the trace monoïd as in Theorem 3.

EXAMPLE (NON-TRANSITIVE CASE). With the same alphabet, let $\theta = \{(a,b), (b,a), (b,c), (c,b)\}$. The θ-descents of a word correspond to factors ba and cb. Then $\gg_\theta$ is not a transitive relation. Observe that the word cba in $I^+ = \{a, b, c, ba, cb, cba\}$ is not a θ-trivial trace. Also, the class of cba is $\{cba, cab, bca\}$ and contains two words having no θ-descents (or no $\gg_\theta$-adjacencies). Nevertheless, (2) implies

$$\sum_{w \in X^*} y^{\theta \operatorname{des}(w)} w = \frac{1}{1 - (a+b+c) + (1-y)(ba+cb) - (1-y)^2 cba}.$$

4. Adjacencies in words

Let X be an alphabet. From X, we construct the *adjacency* alphabet $A = \{a_{xy} : (x, y) \in X \times X\}$. The *adjacency monomial* and the *sieve polynomial* for $w = x_1 x_2 \cdots x_n \in X^*$ of length $n \geq 2$ are defined respectively as $a(w) = a_{x_1x_2} a_{x_2x_3} \cdots a_{x_{n-1}x_n}$ and $\overline{a}(w) = (a_{x_1x_2} - 1)(a_{x_2x_3} - 1) \cdots (a_{x_{n-1}x_n} - 1)$. For $0 \leq n \leq 1$, we set $a(w) = \overline{a}(w) = 1$. In $\mathbb{Z}[A]\langle\langle X \rangle\rangle$, the algebra of formal power series of words in X^* with polynomial coefficients, the following inversion formula holds:

THEOREM 5 ([10, 14, 17, 19]). *According to the adjacencies, the words in X^* are generated by*

$$\sum_{w \in X^*} a(w)w = \frac{1}{1 - \sum_{w \in X^+} \overline{a}(w)w}. \tag{3}$$

If for $u, v \in X$ we set $a_{uv} = y$ when $x \gg_\theta y$ and $a_{uv} = 1$ otherwise, Theorem 2 can be seen as a corollary of Theorem 5. In passing, we mention that J. Hutchinson and H. Wilf [15] have given a closed formula for counting words by adjacencies.

EXAMPLE. The applications we give rely on the fact that setting $a_{xy} = 1$ eliminates all words containing xy as a factor from the right-hand side of (3). Suppose that $X = \{a, b, c\}$. Setting $a_{aa} = r$, $a_{ab} = s$, $a_{ac} = t$ and the remaining $a_{ij} = 1$ in Theorem 5 yields

$$\sum_{w \in X^*} a(w)w = \frac{1}{1 - \sum_{w \in B} \overline{a}(w)w},$$

where $B = \{a^n \mid n \geq 1\} \cup \{a^n b \mid n \geq 0\} \cup \{a^n c \mid n \geq 0\}$. Thus

$$\begin{aligned}
&\sum_{w \in X^*} a(w)w \\
&\quad = \left[1 - \sum_{n \geq 1} (r-1)^{n-1} a^n - b - \sum_{n \geq 1} (r-1)^{n-1}(s-1) a^n b - c - \sum_{n \geq 1} (r-1)^{n-1}(t-1) a^n c\right]^{-1} \\
&\quad = \frac{1 + a - ra}{1 - ra - b - c + (r-s)ab + (r-t)ac}.
\end{aligned}$$

5. The insertion-shift bijection

In applying Theorem 5 to the enumeration of permutations, we make repeated use of the insertion-shift bijection [8] that associates a finite sequence of non-negative integers to a pair (σ, λ) where σ is a permutation and λ is a partition.

Let $\mathbb{N}_+^n$ be the set of words of length n in $\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$. The *rise set*, *rise number*, *inversion number*, and *norm* of $w = i_1 i_2 \cdots i_n \in \mathbb{N}_+^n$ are respectively defined by

$$\begin{aligned}
&\operatorname{Ris} w = \{k : 1 \leq k \leq n-1, i_k < i_{k+1}\}, && \operatorname{ris} w = |\operatorname{Ris} w|, \\
&\operatorname{inv} w = |\{(k, m) : 1 \leq k < m \leq n, i_k > i_m\}|, && \|w\| = i_1 + i_2 \cdots + i_n.
\end{aligned}$$

The set of non-decreasing words in $\mathbb{N}_+^n$ (i.e., partitions with at most n parts) will be denoted by P_n. We have to construct inductively a bijection $f_n : \mathbb{N}_+^n \to S_n \times P_n$. If $n = 1$, then the map $i \mapsto (1, i)$ does the job. Suppose f_{n-1} exists and let $w = i_1 i_2 \cdots i_n \in \mathbb{N}_+^n$. Applying f_{n-1} to the first $n-1$ letters of w gives a pair (α, δ) in $S_{n-1} \times P_{n-1}$. Note $\delta_0 = 0$, $\delta_n = \infty$ and $\delta = (\delta_1 \cdots \delta_{n-1})$. Determining k such that $\delta_{k-1} \leq i_n < \delta_k$, we define $f_n(w)$ to be the pair

$$\Big(\alpha(1) \cdots \alpha(k-1)\, n\, \alpha(k) \cdots \alpha(n-1), \delta_1 \cdots \delta_{k-1}\, i_n\, (\delta_k - 1) \cdots (\delta_{n-1} - 1)\Big).$$

LEMMA 1 ([8]). *For $n \geq 1$ and for $w \in \mathbb{N}_+^n$, if w is mapped to the pair (σ, λ) by the bijection $f_n : \mathbb{N}_+^n \to S_n \times P_n$, then* $\operatorname{Ris} w = \operatorname{Ris} \sigma^{-1}$ *and* $\|w\| = \operatorname{inv} \sigma + \|\lambda\|$.

EXAMPLE. The word $w = 372314 \in \mathbb{N}_+^6$ is mapped by f_6 to the pair $(\sigma, \lambda) = (531426, 111244) \in S_6 \times P_6$. Noting that $\sigma^{-1} = 352416$, we see that $\operatorname{Ris} w = \{1, 3, 5\} = \operatorname{Ris} \sigma^{-1}$ and that $\|w\| = 20 = \operatorname{inv} \sigma + \|\lambda\| = 7 + 13$.

6. q-Eulerian polynomials and bibasic Bessel functions

As the first application of Theorem 1, we derive a generating function for the sequence

$$A_n(t, q) = \sum_{\sigma \in S_n} t^{\operatorname{ris} \sigma} q^{\operatorname{inv} \sigma}.$$

The polynomial $A_n(t, 1)$ is the n-th *Eulerian polynomial*.

We set $a_{ij} = t$ if $i \leq j$ and $a_{ij} = 1$ otherwise. Theorem 1 reduces to

$$\sum_{w \in \mathbb{N}_+^*} t^{\operatorname{ris} w} w = \frac{1}{1 - \sum_{n \geq 1} (t-1)^{n-1} \sum_{i_1 \leq i_2 \leq \cdots \leq i_n} i_1 i_2 \cdots i_n}. \tag{4}$$

Using (4) and lemma 1, we have the following form for the generating function:

$$\sum_{n \geq 0} \frac{A_n(t, q) z^n}{(q)_n} = \frac{1}{1 - \sum_{n \geq 1} (t-1)^{n-1} z^n / (q)_n} = \frac{1 - t}{E\left(-z(1-t), q\right) - t}, \tag{5}$$

where $E(z, q) = \sum_{n \geq 0} z^n / (q)_n$ is a well-known q-analog of e^z.

Now we derive a generating function for the sequence

$$B_n(t, q_1, q_2) = \sum_{(\sigma_1, \sigma_2) \in S_n^2} t^{\operatorname{dd}(\sigma_1, \sigma_2)} q_1^{\operatorname{inv} \sigma_1} q_2^{\operatorname{inv} \sigma_2}.$$

We use the alphabet $X' = \{\binom{a}{a'}, (a, a') \in \mathbb{N}_+^2\}$ and for letters $\mathbf{i} = (i_1, i_2)$ and $\mathbf{j} = (j_1, j_2)$, we set $a_{\mathbf{ij}} = t$ if $i_1 \leq i_2$ and $j_1 \leq j_2$, and $a_{\mathbf{ij}} = 1$ otherwise. Repeating (5) with appropriate modifications gives

$$\sum_{n \geq 0} \frac{B_n(t, q_1, q_2) z^n}{(q_1)_n (q_2)_n} = \frac{1}{1 - \sum_{n \geq 1} (t-1)^{n-1} \frac{z^n}{(q_1)_n (q_2)_n}} = \frac{1 - t}{J\left(-z(1-t), q_1, q_2\right) - t},$$

where $J(z, q) = \sum_{n \geq 0} (-1)^n z^n / (q_1)_n (q_2)_n$ is a bibasic Bessel function.

7. q-Euler polynomials

D. André [1] showed that if E_n denotes the number of up-down alternating permutations in S_n (that is, $\sigma \in S_n$ that $\sigma(1) < \sigma(2) > \sigma(3) < \sigma(4) > \cdots$), then

$$\sum_{n \geq 0} E_n \frac{z^n}{n!} = \frac{1 + \sin z}{\cos z}.$$

The number E_n is known as the n-th *Euler number*.

We now apply Theorem 5 to the more general problem of counting the set of *odd-up permutations* $\mathcal{O}_n = \{\sigma \in S_n : \sigma(1) < \sigma(2), \sigma(3) < \sigma(4) \cdots\}$ by inversion number and by the *number of even indexed rises* $\operatorname{ris}_2 \sigma = |\{k \in \operatorname{Ris} \sigma : k \text{ is even}\}|$. We begin by determining a generating function for

$$C_{2n}(t, q) = \sum_{\sigma \in \mathcal{O}_{2n}} t^{\operatorname{ris}_2 \sigma} q^{\operatorname{inv} \sigma}.$$

Note that $C_{2n}(0, 1) = E_{2n}$.

Let $X = \{\mathbf{i} = (i_1, i_2) : i_1, i_2 \in \mathbb{N}_+ \text{ with } i_1 \leq i_2\}$. For letters $\mathbf{i} = (i_1, i_2)$ and $\mathbf{j} = (j_1, j_2)$, we set $a_{\mathbf{ij}} = t$ if $i_2 \leq j_1$ and $a_{\mathbf{ij}} = 1$ otherwise. Then we have

$$\sum_{n \geq 0} C_{2n}(t, q) \frac{z^{2n}}{(q)_{2n}} = \frac{1}{1 - \sum_{n \geq 1} (t-1)^{n-1} \frac{z^{2n}}{(q)_{2n}}}.$$

Bibliography

[1] André (D.). – Sur les permutations alternées. *Journal de Mathématiques Pures et Appliquées*, vol. 7, 1881, pp. 167–184.

[2] Carlitz (L.). – Eulerian numbers and polynomials. *Math. Magazine*, vol. 33, 1959, pp. 247–260.

[3] Carlitz (L.). – A combinatorial property of q-Eulerian numbers. *American Mathematical Monthly*, vol. 82, 1975, pp. 51–54.

[4] Carlitz (L.), Scoville (R.), and Vaughan (T.). – Enumeration of pairs of permutations. *Discrete Mathematics*, vol. 14, 1976, pp. 215–239.

[5] Cartier (P.) and Foata (D.). – *Problèmes combinatoires de commutation et réarrangements.* – Springer-Verlag, Berlin, 1969, *Lecture Notes in Mathematics*, vol. 85.

[6] Diekert (V.). – Transitive orientations, Möbius functions and complete semi-Thue systems for partially commutative monoïds. In *Automata, Languages and Programming*, *Lecture Notes in Computer Science*, vol. 317, pp. 176–187. – 1988.

[7] Diekert (V.). – *Combinatorics on Traces.* – Springer-Verlag, 1990, *Lecture Notes in Computer Science*, vol. 454.

[8] Fédou (Jean-Marc). – Fonctions de Bessel, empilements et tresses. In Leroux (P.) and Reutenauer (C.) (editors), *Séries Formelles et Combinatoire Algébrique*, *Publications du LaCIM*, vol. 11, pp. 189–202. – Université du Québec à Montréal, Montréal, 1992.

[9] Fédou (Jean-Marc). – Combinatorial objects counted by q-Bessel functions. *Reports of Mathematical Physics*, vol. 34, n° 1, 1994.

[10] Fédou (Jean-Marc) and Rawlings (Don). – Adjacencies in words. *Advances in Applied Mathematics*, 1994. – To appear.

[11] Fédou (Jean-Marc) and Rawlings (Don). – Statistics on finite sequences of permutations. *Electronic Journal of Combinatorics*, 1994. – To appear.

[12] Fédou (Jean-Marc) and Rawlings (Don). – Statistics on pairs of permutations. *Discrete Mathematics*, 1994. – To appear.

[13] Gessel (I. M.). – *Generating Functions and Enumeration of Sequences.* – PhD thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1977.

[14] Goulden (Ian P.) and Jackson (David M.). – *Combinatorial Enumeration.* – John Wiley, New York, 1983.

[15] Hutchinson (J. P.) and Wilf (H. S.). – On Eulerian circuits and words with prescribed adjacency patterns. *Journal of Combinatorial Theory, Series A*, vol. 18, 1975, pp. 80–87.

[16] Riordan (J.). – *An introduction to combinatorial theory.* – John Wiley, New York, 1958.

[17] Stanley (R. P.). – Binomial posets, Möbius inversions, and permutation enumeration. *Journal of Combinatorial Theory, Series A*, vol. 20, 1976, pp. 336–356.

[18] Viennot (X. G.). – Heaps of pieces I: Basic definitions and combinatorial lemmas. In *Combinatoire énumérative. Lecture Notes in Mathematics*, pp. 321–350. – Springer-Verlag, 1986.

[19] Zeilberger (D.). – Enumeration of words by their number of mistakes. *Discrete Mathematics*, vol. 34, 1981, pp. 89–92.

Eulerian Calculus and Transformations of Rearrangements

Dominique Foata
Université Louis Pasteur, Strasbourg

June 13, 1994

[summary by Dominique Gouyou-Beauchamps]

1. Introduction

The purpose of this talk is to study the behaviour of several classical statistics on words, such as the number of descents, the number of excedances, the major index, when the strict inequalities required in their definitions are relaxed to include some equalities[1]. Those new statistics are defined on classes of words with repetitions.

Let X^* be the free monoid generated by a totally ordered alphabet X that for convenience we take as being the subset $[r] = \{1, 2, \ldots, r\}$ $(r \geq 1)$ of the positive integers. X is a fixed non-empty set on which a total ordering D is defined. If $x, y \in X$ and $(x, y) \in D$ we write $x <_D y$ or simply $x < y$ if no confusion can arise. D need not be the standard ordering on $[r]$. We also have a fixed integer, k, such that $0 \leq k \leq r$ and $j = r - k$. The letters $1, \ldots, j$ will be called *small* and the letters $j+1, \ldots, r$ *large*. We say that D is *compatible* with k if, for all x large and y small, we have $x > y$. We also introduce a small letter $*$ which is greater than any small letter of X.

A word w' is said to be a *rearrangement* of the word $w = x_1 x_2 \cdots x_m$ if it can be obtained from w by permuting the letters $x_1, x_2, \ldots, x_m$ in some order. The set of all the rearrangements of a word w will be denoted $C(w)$. Such a set necessarily contains a unique word $v = y_1 y_2 \cdots y_m$ whose letter are in non-decreasing order: $y_1 \leq y_2 \leq \cdots \leq y_m$. It will be convenient to denote $\overline{w}$ the unique non-decreasing word in the class $C(w)$.

Let $w = x_1 x_2 \cdots x_m$ be a word and let $\overline{w} = v = y_1 y_2 \cdots y_m$ be its non-decreasing rearrangement. The *number of excedances*, $\operatorname{exc} w$, and the *number of descents*, $\operatorname{des} w$, of the word w are classically defined as

$$\operatorname{exc} w = \#\{i : 1 \leq i \leq m, x_i > y_i\},$$
$$\operatorname{des} w = \#\{i : 1 \leq i \leq m-1, x_i > x_{i+1}\},$$

while the *major index*, $\operatorname{maj} w$, is the sum of the i's such that $1 \leq i \leq m-1$ and $x_i > x_{i+1}$.

MacMahon [8, p. 186] proved that for each rearrangement class $C(v)$ and each integer j there are as many words $w \in C(v)$ such that $\operatorname{exc} w = j$ as there are words $w' \in C(v)$ such that $\operatorname{des} w' = j$.

Let $\mathbf{c} = (c_1, \ldots, c_j)$ and $\mathbf{d} = (d_1, \ldots, d_k)$ be two vectors with positive integer components. Also let $c = c_1 + \cdots + c_j$, $d = d_1 + \cdots + d_k$ and $c + d = m$. The class of all $m!/(c_1! \cdots c_j! d_1! \cdots d_k!)$ rearrangements of the word $1^{c_1} \cdots j^{c_j} (j+1)^{d_1} \cdots r^{d_k}$ will be denoted by $R(\mathbf{c}, \mathbf{d})$ or by $C(v)$ where v is a given word in $R(\mathbf{c}, \mathbf{d})$.

Let $w = x_1 x_2 \cdots x_m$ be a word and let $\overline{w} = v = y_1 y_2 \cdots y_m$ be its non-decreasing rearrangement (with respect to a given ordering D). We say that the word w has a *k-excedance* at i $(1 \leq i \leq m)$,

[1] The original articles by J. Clarke and D. Foata can be found in [2, 3, 4].

if either $x_i > y_i$, or $x_i = y_i$ and x_i large. We also say that w has a *k-descent* at i $(1 \le i \le m)$, if either $x_i > x_{i+1}$, or $x_i = x_{i+1}$ and x_i large (by convention, $x_{m+1} = *$). The number of k-excedances and k-descents of a word w are denoted by $\mathrm{exc}_k\, w$ and $\mathrm{des}_k\, w$. The *k-major index*, $\mathrm{maj}_k\, w$, is the sum of all i's $(1 \le i \le m)$ such that i is a k-descent.

In [2] R. J. Clarke and D. Foata showed that for each ordering D compatible with $k \ge 0$ the statistics "des_k" and "exc_k" were equidistributed on each rearrangement class $R(\mathbf{c}, \mathbf{d})$. Actually, they constructed a bijection Φ_k of each rearrangement class $R(\mathbf{c}, \mathbf{d})$ onto itself that satisfied $\mathrm{des}_k\, w = \mathrm{exc}_k\, \Phi_k(w)$, identically. Hence for each rearrangement class $R(\mathbf{c}, \mathbf{d})$ the generating polynomials $\sum_w t^{\mathrm{des}_k\, w}$ and $\sum_w t^{\mathrm{exc}_k\, w}$ $(w \in R(\mathbf{c}, \mathbf{d}))$ are equal. Let $A_{\mathbf{c},\mathbf{d}}(t)$ be their common value. It was also shown that the generating function for those polynomials could be expressed as

$$\sum_{\mathbf{c},\mathbf{d}} \frac{\mathbf{u}^{\mathbf{c}}\mathbf{v}^{\mathbf{d}}}{(1-t)^{c+d+1}} A_{\mathbf{c},\mathbf{d}}(t) = \sum_{s \ge 0} t^s \frac{(1+v_1)^s \cdots (1+v_k)^s}{(1-u_1)^{s+1} \cdots (1-u_j)^{s+1}}, \tag{1}$$

where $\mathbf{u}^{\mathbf{c}} = u_1^{c_1} \cdots u_j^{c_j}$ and $\mathbf{v}^{\mathbf{d}} = v_1^{d_1} \cdots v_k^{d_k}$.

As usual, let $(a;q)_n$ denote the q-ascending factorial

$$(a;q)_n = \begin{cases} 1 & \text{if } n = 0, \\ (1-a)(1-aq)\cdots(1-aq^{n-1}) & \text{if } n \ge 1. \end{cases}$$

Then, a natural q-analogue of (1) can read

$$\sum_{\mathbf{c},\mathbf{d}} \frac{\mathbf{u}^{\mathbf{c}}\mathbf{v}^{\mathbf{d}}}{(t;q)_{c+d+1}} A_{\mathbf{c},\mathbf{d}}(t,q) = \sum_{s \ge 0} t^s \frac{(-qv_1;q)_s \cdots (-qv_k;q)_s}{(u_1;q)_{s+1} \cdots (u_j;q)_{s+1}}. \tag{2}$$

The motivation of the talk is to extend Han's construction [6] to weighted words. This consists, first, in finding an appropriate extension "den_k" of the Denert statistic "den" [5], defined by Han, then, of constructing an explicit bijection ρ of each rearrangement class $R(\mathbf{c}, \mathbf{d})$ onto itself such that the equality over the bivariate statistics $(\mathrm{des}_k, \mathrm{maj}_k)(w) = (\mathrm{exc}_k, \mathrm{den}_k)\rho(w)$ holds identically.

2. The "den_k" statistic

Let $S = \{1, \ldots, j\}$ be the set of small elements of X and let $L = \{j+1, \ldots, r\}$ be the set of large elements of X. Let $s_{\max}$ be the largest small letter of X (under the ordering D). Besides the small letter $*$ that satisfies $s_{\max} <_D * <_D b$ for any letter b greater than $s_{\max}$, we also adjoin to X a large letter ∞ that is greater than every letter of X. Define X^+ to be $X \cup \{*, \infty\}$. Similarly, $L^+ = L \cup \{\infty\}$ and $S^+ = S \cup \{*\}$.

Let a and b be elements of X^+. Then we define the *cyclic interval* $]\!]a, b]\!]$ by

$$]\!]a, b]\!] = \begin{cases} (a, b] & \text{if } a \le b, \\ X^+ \backslash (b, a] & \text{otherwise.} \end{cases} \tag{3}$$

Thus $]\!]a, a]\!] = \emptyset$. Further, we define $]\!]a, b]\!]_k$ by

$$]\!]a, b]\!]_k = \begin{cases}]\!]a, b]\!] & \text{if } a, b \in S^+, \\ ]\!]a, b]\!] \cup a & \text{if } a \in L^+, b \in S^+, \\ ]\!]a, b]\!] \setminus b & \text{if } a \in S^+, b \in L^+, \\ ]\!]a, b]\!] \cup a \setminus b & \text{if } a, b \in L^+, a \ne b, \\ X^+ & \text{if } a = b \in L^+. \end{cases} \tag{4}$$

The elements of X^+ can be visualized as points on a circle (or a square!) as shown on Fig. 1. The k-cyclic intervals $[\![a,b]\!]_k$ must be read counterclockwise. The path $\dashv\!\rightarrow$ on the S^+-part shows that whenever a is small, the interval $[\![a,b]\!]_k$ is of the form "$(a,\ldots$" (or "$]a,\ldots$" in the French notation) so that $a \notin [\![a,b]\!]_k$. On the contrary, the path $-\!\!\vdash\!\!\rightarrow$ on the L^+-part shows that $[\![a,b]\!]_k = [a,\ldots$ and so $a \in [\![a,b]\!]_k$ whenever a is large; but $[\![a,b]\!]_k = \ldots,b)$ (or $\ldots,b[$) whenever b is large and $b \notin [\![a,b]\!]_k$. When D is compatible with k, the small letters lie between ∞ and $*$, and the large ones between $*$ and ∞, still reading the square counterclockwise; also $[\![*,\infty]\!]_k = L$.

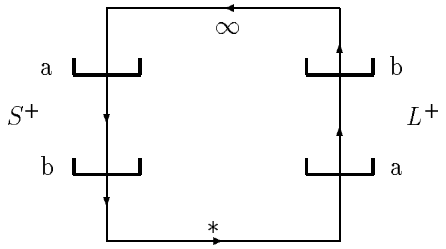

Fig. 1

Let $w = x_1x_2\cdots x_m$ be a word on the letters in X. Put $x_{m+1} = *$, $x_{m+2} = \infty$. For $i = 1,\ldots,m+2$ let $\mathrm{Fact}_i w$ be the left factor $x_1x_2\cdots x_{i-1}$ of w and for each subset B of X let $\mathrm{Fact}_i w \cap B$ be the subword of $\mathrm{Fact}_i w$ consisting only of those letters of $\mathrm{Fact}_i w$ that are in B. Furthermore, let $|\mathrm{Fact}_i w \cap B|$ denote the length of that subword.

Now let $\overline{w} = y_1y_2\cdots y_m$ the non-decreasing rearrangement of a word $w = x_1x_2\cdots x_m$. The den_k-coding of w is defined to be the sequence $(s_i)_{1\le i\le m+1}$, where

$$s_i = \begin{cases} |\,\mathrm{Fact}_i\, w \cap [\![x_i,y_i]\!]_k\,| & \text{if } 1 \le i \le m, \\ |w \cap L| & \text{if } i = m+1, \end{cases} \tag{5}$$

and the statistic $\mathrm{den}_k\, w$ to be

$$\mathrm{den}_k\, w = \sum_{i=1}^{m+1} s_i. \tag{6}$$

THEOREM 1. *Let v be a fixed word in X^* and let D and E be total orderings on $X = [r]$. Assume that both D and E are compatible with k. Then there is a bijection μ on $C(v) = R(\mathbf{c},\mathbf{d})$ onto itself such that for all $w \in C(v)$,*

$$(\mathrm{des}_{k,D}, \mathrm{maj}_{k,D})w = (\mathrm{des}_{k,E}, \mathrm{maj}_{k,E})\mu(w).$$

THEOREM 2. *Let v be a fixed word in X^* and let D be an ordering on X compatible with k. Then there is a bijection ρ on $C(v)$ onto itself such that for all $w \in C(v)$,*

$$(\mathrm{des}_k, \mathrm{maj}_k)w = (\mathrm{exc}_k, \mathrm{den}_k)\rho(w).$$

THEOREM 3. *Let v be a fixed word in X^* and let D and E be total orderings on $X = [r]$. Then there is a bijection δ on $C(v) = R(\mathbf{c},\mathbf{d})$ onto itself such that for all $w \in C(v)$,*

$$(\mathrm{exc}_{k,D}, \mathrm{den}_{k,D})w = (\mathrm{exc}_{k,E}, \mathrm{den}_{k,E})\delta(w).$$

Now we calculate the distribution of $(\mathrm{des}_k, \mathrm{maj}_k)$. Let

$$A_{\mathbf{c},\mathbf{d}}(t,q) = \sum_w t^{\mathrm{des}_k\, w} q^{\mathrm{maj}_k\, w} \qquad (w \in R(\mathbf{c},\mathbf{d}))$$

be the generating function for the pair $(\mathrm{des}_k, \mathrm{maj}_k)$ over the class $R(\mathbf{c}, \mathbf{d})$.

THEOREM 4. *The factorial generating function for the polynomials $A_{\mathbf{c},\mathbf{d}}(t,q)$ satisfies (2).*

3. The Han transposition

In this section we describe the "Han transposition", a way of manipulating biwords that preserves the statistics "den" and "exc".

A *biword* is a two rowed matrix $\alpha = \binom{u}{w}$, where u and v are words in X^* of the same length. The biword α is called a *circuit* if u is a rearrangement of w. A circuit $\alpha = \binom{y_1 y_2 \cdots y_m}{x_1 x_2 \cdots x_m}$ is called a *cycle*, if $y_m = x_1$ and $y_i = x_{i+1}$ for $i = 1, \ldots, m-1$.

Let $x, y, a, b \in X \cup \{*\}$. Then a and b are *neighbours* with respect to (x,y) if both a and b are in $]\!]a,b[\![_k$ or neither in $]\!]a,b[\![_k$. Otherwise, a and b are strangers with respect to (x,y).

Consider the biword $\binom{u}{w}$, where $u = y_1 y_2 \cdots y_m$ and $w = x_1 x_2 \cdots x_m$. An ordering D of X being given, we define

$$\mathrm{exc}_k \binom{u}{w} = |\{i : 1 \le i \le m \quad \text{and} \quad x_i > y_i \quad \text{or} \quad x_i = y_i \in L^+\}|;$$

$$\mathrm{den}_k \binom{u}{w} = \sum_{i=1}^{m} |\mathrm{Fact}_i w \cap]\!]x_i, y_i[\![_k|.$$

If $\overline{w}$ is the non-decreasing rearrangement of w, then clearly $\mathrm{exc}_k \binom{\overline{w}}{w} = \mathrm{exc}_k\, w$; and by (5) and (6) $\mathrm{den}_k\, w = \mathrm{den}_k \binom{\overline{w}}{w} + |w \cap L|$. Note that if D is compatible with k, $\mathrm{den}_k\, w = \mathrm{den}_k \binom{\overline{w}\infty}{w*}$.

Let $\binom{xy}{ab}$ be a biword of length two. Following Han [6] we define the Han transposition T by

$$T\binom{xy}{ab} = \begin{cases} \binom{yx}{ab} & \text{if } a \text{ and } b \text{ are neighbours,} \\ \binom{yx}{ba} & \text{if } a \text{ and } b \text{ are strangers.} \end{cases} \tag{7}$$

If $\alpha = \binom{u}{w} = \binom{y_1 y_2 \cdots y_m}{x_1 x_2 \cdots x_m}$ is a biword of length m and $1 \le i < m$, we define $T_i\alpha$ to be the biword obtained when the biword $\beta = \binom{y_i y_{i+1}}{x_i x_{i+1}}$ consisting of the i-th and $(i+1)$th columns of α is replaced by $T\beta$.

LEMMA 1. *Let $\alpha = \binom{u}{w}$ be a biword of length m and let $1 \le i < m$. Then*

$$(\mathrm{exc}_k, \mathrm{den}_k) T_i \alpha = (\mathrm{exc}_k, \mathrm{den}_k)\alpha.$$

Let z_1, z_2 be two distinct letters of $X \cup \{*\}$ and v be a word of length $m-1$ in the alphabet $X \cup \{*\} \setminus \{z_2\}$. Denote by $\mathcal{C}(v, z_1, z_2)$ the set of all biwords $\alpha = \binom{u}{w}$, where u is the non-decreasing rearrangement of vz_2 and w is any rearrangement of vz_1. Thus u has one occurrence of z_2, while w has none. However the occurrences of the other letters are the same, except for z_1 that occurs one more time in w than in u.

If z_2 is the i-th letter in the word u, the product $T_{m-1} \cdots T_{i+1} T_i$ will transform α into a biword of the form $\alpha' = \binom{u' z_2}{w' y_1}$. Then, either u' has no occurrence of y_1, in which case $y_1 = z_1$ and w' must be a rearrangement of u', or y_1 does occur in u'. In the former case, define $T_{z_2}(\alpha) = \binom{u' z_2}{w' z_1}$. In the latter case, the rightmost occurrence of y_1 in u' is, say, its i'-th letter. Then the product $T_{m-2} \cdots T_{i'+1} T_{i'}$ transforms α' into a biword of the form $\alpha'' = \binom{u'' y_1 z_2}{w'' y_2 y_1}$. Again, either u'' has no occurrence of y_2, in which case $y_2 = z_1$ and w'' must be a rearrangement of u'', or y_2 does occur in u''. In the former case, define $T_{z_2}(\alpha) = \binom{u'' y_1 z_2}{w'' z_1 y_1}$. In the latter case, we continue the same procedure as before by moving the rightmost occurrence of y_2 in u'' to the right of u''. After finitely many

steps we reach a biword $\alpha^{(l)} = \binom{u^{(l)}y_{l-1}\cdots y_1z_2}{w^{(l)}y_l\cdots y_2y_1}$, where $u^{(l)}$ has no occurrence of y_l. Then necessarily $y_l = z_1$ and $w^{(l)}$ is a rearrangement of $u^{(l)}$. Note that $u^{(l)}$ may be empty. Define $u_1 = u^{(l)}$, $w_1 = w^{(l)}$, $v_1 = y_{l-1}\cdots y_2y_1$ and

$$T_{z_2}(\alpha) = \begin{pmatrix} u^{(l)}y_{l-1}\cdots y_1z_2 \\ w^{(l)}y_l\cdots y_2y_1 \end{pmatrix} = \begin{pmatrix} u_1v_1z_2 \\ w_1z_1v_1 \end{pmatrix}. \tag{8}$$

Thus for each α in $\mathcal{C}(v, z_1, z_2)$ there is a well-defined product of Han transpositions that maps α onto a biword of the form $\binom{u_1v_1z_2}{w_1z_1v_1}$, where u_1 is the non-decreasing rearrangement of w_1 with no occurrence of z_1. Denote by $\mathcal{D}(v, z_1, z_2)$ the set of biwords of the previous form $\binom{u_1v_1z_2}{w_1z_1v_1}$.

LEMMA 2. *The mapping $T_{z_2} : \mathcal{C}(v, z_1, z_2) \to \mathcal{D}(v, z_1, z_2)$ is a bijection.*

The inverse mapping applied to the biword $\binom{u_1v_1z_2}{w_1z_1v_1}$ in $\mathcal{D}(v, z_1, z_2)$ is derived by moving to the left, to the first position where the resulting word is non-decreasing, successively the first, the second, ..., the last letter z_2 of the word v_1z_2, the move being made by means of the Han transpositions as defined in (7). The inverse mapping is then independent of z_2 and will be denoted by T^{-1}.

4. The den-maj bijection

In this section we give the main tools for proving Theorem 2. We assume that the ordering D is compatible with k, and in fact that D is the standard ordering on $X = [r]$. If y is a letter and w a non-decreasing word, we will write $y < w$ (resp. $y \le w$), if y is less than (resp. less than or equal to) all the letters in w.

As for the first fundamental transformation described in Cartier and Foata [1] or in Lothaire [7, chap. 10] and Han's fundamental bijection [6] we need an appropriate *word factorization*. That factorization can be built as follows.

Let $z \in X \cup \{*\}$ and let v be a word in that alphabet, then the word zv is said to be *k-dominant*, if z is large and all letters in v are greater than or equal to z, or if z is small and small letters in v are less than or equal to z.

Every word in the alphabet $X \cup \{*\}$ has a unique factorization $(z_1v_1, z_2v_2, \ldots, z_nv_n)$ (the z_i's are letters and the v_i's words), called its *k-factorization*, having the following properties:

(1) $w = z_1v_1z_2v_2\cdots z_nv_n$;
(2) each factor z_iv_i is k-dominant $(1 \le i \le n)$;
(3) there exists an integer l $(1 \le l \le n)$ such that $z_1 > z_2 > \cdots > z_l > *$ and $z_{l+1} < z_{l+2} < \cdots < z_n \le *$.

The *k-factorization* of a word w may be obtained as follows: a letter z of w is called a *k-record*, if either z is large and all letters to the left of z are larger than z, or z is small and all small letters to the left of z are smaller than z. The *k-factorization* of w is then obtained by cutting w before each k-record.

The main property of the *k-factorization* on which our transformation is based is the following: let $(z_1v_1, z_2v_2, \ldots, z_nv_n)$ be the *k-factorization* of a word w. Then for each $i = 1, \ldots, n-1$ no letter in the left factor $(z_1v_1, \ldots, z_iv_i)$ is equal to z_{i+1} or is strictly between z_i and z_{i+1}. Let w_1 be any rearrangement of that factor; then the *k-factorization* of $(w_1z_{i+1}v_{i+1}\cdots z_nv_n)$ has the same rightmost $(n-i)$ factors $(z_{i+1}v_{i+1}, \ldots, z_nv_n)$ as w and the same rightmost $(n-i+1)$ *k-records* $z_i, z_{i+1}, \ldots, z_n$ as w.

Consider a biword $\alpha = \binom{u_1u_2\infty}{w_1zu_2}$, where:

(1) u_1, u_2, w_1 are words in the alphabet $X \cup \{*\}$;
(2) u_1 is the non-decreasing rearrangement of w_1;

(3) z is a k-record of the word $w_1 z u_2$.

Such a biword is called a *supercycle*. A supercycle is said to be *initial*, if u_1 and w_1 are empty. The notion of *final* supercycle will be defined shortly.

LEMMA 3. *If α is a supercycle, the factorization*

$$\alpha = \left(\begin{array}{c|cc} u_1 & u_2 & \infty \\ w_1 & z & u_2 \end{array} \right), \tag{9}$$

where u_1 is the non-decreasing rearrangement of some left factor of the bottom word of α and where z is a k-record of the bottom word, is unique. The factorization (9) is called the canonical form of α.

Let α be a supercycle written in its canonical form as in (9) and let $(z_1 v_1, z_2 v_2, \ldots, z_n v_n)$ be the *k-factorizarion* of $z u_2$. Then the following factorization of α, indicated by vertical bars,

$$\alpha = \left(\begin{array}{c|cc|cc|c|cc} u_1 & v_1 & z_2 & v_2 & z_3 & \cdots & v_n & \infty \\ w_1 & z_1 & v_1 & z_2 & v_2 & \cdots & z_n & v_n \end{array} \right) \tag{10}$$

is well defined. Call it the *k-factorization* of the supercycle α. The positive integer n, which is the number of factors in the k-factorization of $z u_2$, is called the *index* of α and denoted by $\mathrm{index}(\alpha)$. A supercycle α is said *final*, if its index is equal to 1.

Let α be a supercycle as shown in (10), supposed to be non final, so that $n \geq 2$. With the notation of (8) the left factor $\binom{u_1 v_1 z_2}{w_1 z_1 v_1}$ of the supercycle α is an element of $\mathcal{D}(u_1 v_1, z_1, z_2)$. Apply the inverse transformation T^{-1} to that left factor. We get a biword $\binom{u_1''}{w_1'} \in \mathcal{C}(u_1 v_1, z_1, z_2)$. Then form the supercycle

$$\alpha'' = \left(\begin{array}{c|cc|c|c} u_1'' & v_2 & z_3 & \cdots & v_n & \infty \\ w_1' & z_2 & v_2 & \cdots & z_n & v_n \end{array} \right). \tag{11}$$

Replacing the only occurrence of z_2 in u_1'' by z_1 transforms u_1'' into a true rearrangement u_1' of w_1'. Furthermore, u_1' is non-decreasing and we then obtain a supercycle

$$\alpha' = \left(\begin{array}{c|cc|c|c} u_1' & v_2 & z_3 & \cdots & v_n & \infty \\ w_1' & z_2 & v_2 & \cdots & z_n & v_n \end{array} \right). \tag{12}$$

Moreover, the above expression derived from the k-factorization of α is precisely the k-factorization of α' and the rightmost k-record of w_1' is equal to z_1. Finally, $\mathrm{index}(\alpha') = n - 1$. Thus the mapping $\tau : \alpha \mapsto \alpha'$ is well defined and satisfies

$$\mathrm{index}(\tau(\alpha)) < \mathrm{index}(\alpha), \tag{13}$$

if α is not final.

LEMMA 4. *Let v a non-decreasing word in the alphabet $X \cup \{*\}$ and let $S(v)$ be the set of the supercycles $\alpha = \binom{u_1 u_2 \infty}{w_1 z u_2}$, whose bottom word $w_1 z u_2$ is a rearrangement of v. If α is not initial, there is a unique $\beta \in S(v)$ such that $\tau(\beta) = \alpha$ and index(α) <index(β).*

LEMMA 5. *For each supercycle α which is not final, we have*

$$(\mathrm{exc}_k, \mathrm{den}_k)\tau(\alpha) = (\mathrm{exc}_k, \mathrm{den}_k)\alpha.$$

LEMMA 6. *If w is a word in the alphabet X and $\alpha = \binom{u * \infty}{w *}$ is an initial supercycle, then*

$$(\mathrm{exc}_k, \mathrm{den}_k)\alpha = (\mathrm{des}_k, \mathrm{maj}_k)\alpha = (\mathrm{des}_k, \mathrm{maj}_k)w.$$

The bijection of Theorem 2 is constructed as follows:

(1) Let $w = x_1 x_2 \cdots x_m = x_1 u$ be a word in the alphabet $X = [r]$; form the initial supercycle $\alpha = \binom{u * \infty}{w *}$;
(2) Apply the mapping τ to α iteratively until a final supercycle is reached. This makes sense because of (13). Furthermore, when applying τ iteratively, the letter $*$ remains the rightmost letter in all the supercycles within the iteration. Denote by $\tilde{\alpha} = \binom{\overline{w}\infty}{\tilde{w}*}$ the final supercycle obtained. Then $\overline{w}$ is the non-decreasing rearrangement of w and $\tilde{w}$;
(3) Define ρ by $\rho(w) = \tilde{w}$.

Theorem 2 follows from lemma (5) and lemma (6).

EXAMPLE. Consider the order $1 < 2 < 3 < * < 4 < 5 < \infty$ ($k = 2$ and $4, 5$ large) and start with the word $w = 4, 4, 5, 1, 3, 1, 2, 3, 5$, so that (indicating k-factorization by vertical bars) the initial supercycle is

$$\alpha = \left(\begin{array}{c|c|c|c} 4\,5\,1 & 3 & 1\,2\,3\,5\,* & \infty \\ 4\,4\,5 & 1 & 3\,1\,2\,3\,5 & * \end{array} \right)$$

First, $T^{-1}\binom{4\,5\,1}{4\,4\,5} = T_1 T_2 \binom{4\,5\,1}{4\,4\,5} = \binom{1\,4\,5}{4\,5\,4}$, so that

$$\alpha'' = \left(\begin{array}{c|c|c|c} 1\,4\,5 & 3 & 1\,2\,3\,5\,* & \infty \\ 4\,5\,4 & 1 & 3\,1\,2\,3\,5 & * \end{array} \right)$$

(in the notation of (11)). To obtain $\alpha' = \tau(\alpha)$ we have to replace $z_2 = 1$ by $z_1 = 4$, so that

$$\alpha_2 = \alpha' = \left(\begin{array}{c|c|c|c} 4\,4\,5 & 3 & 1\,2\,3\,5\,* & \infty \\ 4\,5\,4 & 1 & 3\,1\,2\,3\,5 & * \end{array} \right).$$

Next $T_{-1}\binom{4\,4\,5\,3}{4\,5\,4\,1} = T_1 T_2 T_3 \binom{4\,4\,5\,3}{4\,5\,4\,1} = \binom{3\,4\,4\,5}{4\,5\,1\,4}$ and

$$\alpha'' = \left(\begin{array}{c|c|c} 3\,4\,4\,5 & 1\,2\,3\,5\,* & \infty \\ 4\,5\,1\,4 & 3\,1\,2\,3\,5 & * \end{array} \right).$$

To get the next supercycle we have to replace $z_2 = 3$ by $z_1 = 1$, so that

$$\alpha_3 = \left(\begin{array}{c|c|c} 1\,4\,4\,5 & 1\,2\,3\,5\,* & \infty \\ 4\,5\,1\,4 & 3\,1\,2\,3\,5 & * \end{array} \right).$$

Next the transformation T^{-1} to be applied to $\binom{1\,4\,4\,5\,1\,2\,3\,5\,*}{4\,5\,1\,4\,3\,1\,2\,3\,5}$ is $(T_5T_6T_7T_8)(T_4T_5T_6)(T_3T_4T_5)(T_2T_3T_4)$, as we have to move the second 1, 2, 3 and $*$ to the left. We then get $T^{-1}\binom{1\,4\,4\,5\,1\,2\,3\,5\,*}{4\,5\,1\,4\,3\,1\,2\,3\,5} = \binom{1\,1\,2\,3\,*\,4\,4\,5\,5}{4\,5\,1\,4\,1\,3\,2\,3\,5}$, so that

$$\alpha'' = \left(\begin{array}{c|c} 1\,1\,2\,3\,*\,4\,4\,5\,5 & \infty \\ 4\,5\,1\,4\,1\,3\,2\,3\,5 & * \end{array} \right).$$

Finally, $*$ on the top row is to be replaced by the penultimate k-record, i.e., 3. We get

$$\tilde{\alpha} = \left(\begin{array}{c|c} 1\,1\,2\,3\,3\,4\,4\,5\,5 & \infty \\ 4\,5\,1\,4\,1\,3\,2\,3\,5 & * \end{array} \right).$$

Thus $\rho(w) = 4, 5, 1, 4, 1, 3, 2, 3, 5$. We can verify that

$$(\mathrm{des}_k, \mathrm{maj}_k)w = (\mathrm{des}_k, \mathrm{maj}_k)\alpha = (\mathrm{exc}_k, \mathrm{den}_k)\alpha = (\mathrm{exc}_k, \mathrm{den}_k)\tilde{\alpha} = (\mathrm{exc}_k, \mathrm{den}_k)\rho(w) = (4, 18).$$

Bibliography

[1] Cartier (P.) and Foata (D.). – *Problèmes combinatoires de commutation et réarrangements*. – Springer Verlag, Berlin, 1969, *Lecture Notes in Mathematics*, vol. 85.

[2] Clarke (Robert J.) and Foata (Dominique). – Eulerian calculus, I: univariable statistics. – Preprint, 1993.

[3] Clarke (Robert J.) and Foata (Dominique). – Eulerian calculus, II: an extension of Han's fundamental transformation. – Preprint, 1994.

[4] Clarke (Robert J.) and Foata (Dominique). – Eulerian calculus, III: the ubiquitous Cauchy formula. – Preprint, 1994.

[5] Denert (Marlene). – The genus zeta function of hereditary orders in central simple algebras over global fields. *Mathematics of Computation*, vol. 54, 1990, pp. 449–465.

[6] Han (Guo-Niu). – Une transformation fondamentale sur les réarrangements de mots. *Advances in Mathematics*, vol. 105, 1994, pp. 26–41.

[7] Lothaire (M.). – *Combinatorics on Words*. – Addisson-Wesley, Reading, 1983, *Encyclopedia of Mathematics and its Applications*, vol. 17.

[8] MacMahon (P. A.). – *Combinatorial Analysis*. – Cambridge University Press, 1915, vol. 1. (Reprinted by Chelsea, New York, 1955).

Part 2

Symbolic Computation

Linear Differential Equations and Liouvillian Solutions

Felix Ulmer
Université de Rennes I

May 30, 1994

[summary by Jacques-Arthur Weil]

Let k be a differential field (e.g. $k = \mathbb{Q}(x)$ or $k = \mathbb{C}(x)$) with derivation $\frac{d}{dx}$. We review the methods of differential Galois theory used for solving the equation $L(y) = a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_0 y = 0$ (with $a_i \in k$). For effectivity and simplicity, we take $k = \overline{\mathbb{Q}}(x)$ in the sequel.

1. Classes of solutions

1.1. A solution is *rational* if it belongs to k. For example, the equation $a_2 y'' + xy' - y = 0$ has the solution $y = x$ which is in k. Algorithms for computing such solutions have been known for long. The first one is due to Liouville (1833). Some faster or more general versions have been given by Abramov [1], Bronstein [2], and Singer [4] (for the case when k contains a wider class of functions).

If there is no rational solution, then one must perform a field extension to find a solution. Let K be a differential field which is an extension of k, and Δ be the derivation on K (resp. δ on k). We say K is a differential field extension of k if Δ and δ coincide on k.

1.2. A solution y of L is *algebraic* if it belongs to an algebraic extension of k. In other words, there is an irreducible polynomial P with coefficients in k such that $P(y) = 0$. For example, if we define y as a zero of the polynomial $y^2 - x$, then y is a solution of $2xy' = y$. Work on characterising such solutions has been performed for example by Pépin, Klein, Jordan, Fuchs, Baldassari & Dwork, Singer (see e.g. [3, 6, 7] for further references).

1.3. A solution that is not algebraic is transcendental. An interesting class of solutions corresponds to the notion of "integrability by quadratures". A solution y of L is *Liouvillian* if it belongs to a field K such that:

(1) $K = K_n \supseteq \cdots \supseteq K_1 \supseteq K_0 = k$;
(2) $K_i = K_{i-1}(\eta_i)$ for $i = 1, \ldots, n$ and:
  (a) η_i is algebraic over K_{i-1}, or
  (b) $\eta_i' \in K_{i-1}$ (case of an integral), or
  (c) $\eta_i'/\eta_i \in K_{i-1}$ (case of exponential of an integral).

For example, if we take $L(y) = y'' - \frac{1}{2(x+1)} y' - (x+1)y = 0$, then $\{\exp[\int \sqrt{1+x}], \exp[-\int \sqrt{1+x}]\}$ forms a basis of liouvillian solutions.

1.4. There is a very important subclass of the liouvillian solutions: we say that a solution y is *exponential* if its logarithmic derivative is in k, i.e. $y'/y \in k$. For example, the equation $y'' - (2 + 4x^2)y = 0$ has the solution $y = e^{x^2}$ $(y'/y = 2x)$. Methods for computing such solutions have been given, for example, by Singer or Bronstein [2, 4].

2. Differential Galois theory

The main known tool to compute liouvillian solutions of linear differential equations is differential Galois theory. Roughly, the idea is to look at the group of transformations that send a solution of the equation to another solution of the equation; from the knowledge of this group, one can derive algebraic properties of the solutions. We now outline this formalism.

2.1. Picard-Vessiot extensions. To a given vector space of solutions of L, one associates a field extension the following way. Since we work in a differential context, in order to adjoin an element y to k we must also add all its derivatives. We write $k\langle y\rangle := k(y, y', y'', \dots)$. We say that $K \supset k$ is a *Picard-Vessiot extension* if $K = k\langle y_1, \dots, y_n\rangle$, where $\{y_1, \dots, y_n\}$ is a basis of the solution space of $L(y) = 0$, and K and k have the same field of constants C (elements with zero-derivative).

Then, we proceed as in classical Galois theory: The *differential Galois group* of L is the set $\mathrm{Gal}(L)$ of the automorphisms of K that let k point-wise fixed *and that commute with the derivation* (this definition does not depend on K). As in classical Galois theory, an element is in k if and only if it is left fixed by $\mathrm{Gal}(L)$; also, the subfields of K appear as fixed fields of some algebraic subgroup of $\mathrm{Gal}(L)$.

2.2. Galois group. Call V the vector space of solutions of L. As $\mathrm{Gal}(L)$ acts on V, we can decompose its action on a basis of V. The image of a solution of L is still a solution of L, so the image of an element of K is completely characterised by the images of the y_i in the basis $\{y_1, \dots, y_n\}$. This provides a faithful matrix representation of degree n of the Galois group: $\mathrm{Gal}(L)$ can be viewed as a subgroup of $GL(n, C)$ (the group of invertible $n \times n$ matrices with entries in C).

In fact, $\mathrm{Gal}(L)$ is a *linear algebraic group* (its entries are solutions of a set of polynomial equations). So, the group has a structure of an algebraic variety. In particular, there is a component of this variety in which lies the origin; we denote it by $\mathrm{Gal}(L)^\circ$. A key fact is that L has a liouvillian solution if and only if $\mathrm{Gal}(L)^\circ$ is solvable (Picard-Vessiot, Kolchin). In this sense, finding liouvillian solutions is the differential analog of searching for solutions by radicals in the classical case.

2.3. Ricatti equation. A theorem of Lie-Kolchin on triangularization of matrix groups implies that $\mathrm{Gal}(L)^\circ$ is solvable if the elements of $\mathrm{Gal}(L)^\circ$ have a common eigenvector y: $\forall \sigma \in \mathrm{Gal}(L)^\circ, \exists c_\sigma \in C, \sigma(y) = c_\sigma y$. As a consequence $\sigma(\frac{y'}{y}) = \frac{\sigma(y')}{\sigma(y)} = \frac{c_\sigma y'}{c_\sigma y} = \frac{y'}{y}$, which means that y'/y is in the fixed field K° of $\mathrm{Gal}(L)^\circ$. This in turn implies that y'/y is algebraic over k.

As a consequence, there exists a $u = y'/y$ that is a solution of $P(u) = u^N + b_{N-1}u^{N-1} + \cdots + b_0 = 0$ and conversely, $y = \exp[\int u]$ is a solution of $L(y) = 0$. If we let $y' = uy$, then $y^{(i)} = R_i(u, u', \dots)y$, with $R_i = R'_{i-1} + uR_{i-1}$. Replacing in L, we get that $\sum a_i R_i(u, u', \dots) = 0$: this is a non-linear differential equation of order $n - 1$ satisfied by u, called the *Ricatti equation*. For example, if $L = y'' - ry$, then the Ricatti equation is $u' + u^2 - r = 0$.

Finding a liouvillian solution is thus reduced to finding an algebraic solution of the Ricatti equation, which again splits into two subproblems: (1) find a bound for the degree N of P; (2) given N, compute the coefficients of a polynomial P such that its zeroes are logarithmic derivatives of solutions of L.

Problem (1) is solved by group-theoretic considerations. It follows from works of Kovacic or Singer that there is a function $f(n)$ such that $N \leq f(n)$ (e.g., $f(2) = 60, f(3) = 360, f(4) \leq 5040, f(5) \leq 25920, f(6) \leq 604800, \dots$). Recent works of Ulmer and Singer & Ulmer show that sharp bounds are $N \leq 12$ for $n = 2$ and $N \leq 36$ for $n = 3$. We shall come back to this point later and we now focus on the actual computation of the coefficients of the polynomial P.

3. Computing a solution

3.1. Symmetric powers. Suppose for a moment that we work in an algebraic closure of k. There, P has N zeroes $u_1, \ldots, u_N$, and $P(u) = \prod (u - u_i)$. Since all zeroes of P are logarithmic derivatives of solutions of $L(y) = 0$, there are N solutions y_i such that the coefficient b_{N-1} satisfies $b_{N-1} = \frac{y_1'}{y_1} + \cdots + \frac{y_N'}{y_N} = \frac{(y_1 \cdots y_N)'}{y_1 \cdots y_N}$. For any integer m, one can construct a linear differential equation $L^{\circledS m}$, called the *m-th symmetric power of L*, whose solution space is spanned by all monomials of degree m in the $y_1, \ldots, y_n$. In particular, b_{N-1} is the logarithmic derivative of a solution of $L^{\circledS N}$: our problem is now reduced to finding exponential solutions of $L^{\circledS N}$. Similar techniques yield the other coefficients.

3.2. Reducible operators. Let $D = \frac{d}{dx}$. Then, $L(y)$ can be viewed as the action of the operator $\sum a_i D^i$ on y. Such operators form a non-commutative multiplicative ring $\mathcal{D} = k[D]$ in the following way: for $a \in k$, we have $D(ay) = aD(y) + a'y$, which give the multiplication rule on $\mathcal{D}$: $Da = aD + a'$ ($\mathcal{D}$ is called an Ore ring of type "derivation"). Before searching for solutions, one should first search if L factors in $\mathcal{D}$. For example, we have $D^2 = D.D = (D + 1/x)(D - 1/x)$. Algorithms performing such factorizations (or detecting reducibility) exist on $\mathbb{C}(x)$. The classical algorithm dates back to Beke/Schlesinger (1895); Grigor'ev, Singer, or Van Hoeij have recently proposed alternative methods.

In terms of solution space, $\mathrm{Gal}(L)$ has an invariant subspace of dimension m if and only if L has a factor of order m. In that case, we say that $\mathrm{Gal}(L)$ (resp. L) is *reducible*.

3.3. Irreducible operators. Assume that $\mathrm{Gal}(L)$ is irreducible. We say that $\mathrm{Gal}(L)$ is *imprimitive* if V is a direct sum of subspaces that are permuted transitively under the action of $\mathrm{Gal}(L)$. Otherwise, it is *primitive*. In general, if $\mathrm{Gal}(L)$ is irreducible then: either $\mathrm{Gal}(L)$ is imprimitive and $\exists y$ with $[k(y'/y) : k]$ small, or $\mathrm{Gal}(L)$ is primitive finite and $\exists y$ with $[k(y'/y) : k]$ big, or $\mathrm{Gal}(L)$ is primitive infinite and there is no liouvillian solution. This is made precise by the following theorems.

THEOREM 1 (KOVACIC, 1986). *Let L be of order 2 and $\mathrm{Gal}(L) \subseteq SL(2, \mathbb{C})$, then:*

(1) $\mathrm{Gal}(L)$ *is reducible, or*
(2) $\mathrm{Gal}(L)$ *is imprimitive and then $\exists y$ with $[k(y'/y) : k] = 2$, or*
(3) $\mathrm{Gal}(L)$ *is primitive and $\exists y$ with $[k(y'/y) : k] = 4, 6, 12$, or*
(4) $\mathrm{Gal}(L) = SL(2, \mathbb{C})$ *and $L(y) = 0$ has no liouvillian solution.*

THEOREM 2 (SINGER-ULMER, 1993). *Let L be of order 3 and $\mathrm{Gal}(L) \subseteq SL(3, \mathbb{C})$, then:*

(1) $\mathrm{Gal}(L)$ *is reducible and $L = L_1(L_2)$ or*
(2) $\mathrm{Gal}(L)$ *is imprimitive and then $\exists y$ with $[k(y'/y) : k] = 3$, or*
(3) $\mathrm{Gal}(L)$ *is primitive finite and $\exists y$ with $[k(y'/y) : k] = 6, 9, 21, 36$, or*
(4) *Else, $L(y) = 0$ has no liouvillian solutions.*

3.4. Algebraic solutions of L. In general, it is difficult to compute y from the knowledge of y'/y (Abel's problem), but one can compute y directly in the case of a known finite primitive group because y is then algebraic. It follows that y is algebraic over $k(y'/y)$, and one can show that there is an integer m such that $y^m \in k(y'/y)$. Thus, if d is one of the possible degrees for $[k(y'/y) : k]$, the minimum polynomial of y is of the form $P(y) = y^{m.d} + a_{d-1} y^{m.(d-1)} + \cdots + a_1 y^m + a_0$. This polynomial has the same number of coefficients as the minimum polynomial of an algebraic solution of the Ricatti equation. To show that the Ricatti equation had an algebraic solution, we showed that there was a subgroup of L with a common eigenvector. Such a subgroup is called

1-reducible. To find the group or a solution, we must therefore find a 1-reducible subgroup H of $\mathrm{Gal}(L)$ of minimal index. Suppose we have found such an H and let $\mathfrak{S} = \{\sigma_1, \ldots, \sigma_d\}$ be a system of representatives of $\mathrm{Gal}(L)/H$. If y_0 is the eigenvector of H, then

$$P(y) = \prod_{\sigma \in \mathfrak{S}} (y^m - \sigma(y_0)^m).$$

Now, as the a_i are rational, they are invariant under $\mathrm{Gal}(L)$. So, one can decompose the a_i in terms of invariants (or semi-invariants) of the group. Recall that a homogeneous polynomial $M(y_1, \ldots, y_n)$ is called an *invariant* of the group if it is left invariant under the action of the group $(\sigma(M)(y_i) = M(\sigma(y_i)) = M(y_i))$. Now, to detect if the group has invariants of degree m (resp. semi-invariants), one just has to search for rational solutions (resp. exponential solutions) of $L^{\circledS m}$, and we are almost done: as these solutions are given up to multiplication by constants, we just adjust the constants so as to really obtain the desired polynomials. Examples and more precise descriptions of this process are given in [5, 6].

4. Symmetric powers

The whole philosophy was to reduce the computation of Liouvillian solutions to the computation of exponential (and sometimes rational) solutions of some symmetric powers of L. In fact, group-theoretic considerations show that one can reduce the presence of liouvillian solutions to the reducibility of some symmetric powers. Conversely, reducibility of some symmetric powers helps finding the Galois group of a given linear differential equation.

THEOREM 3 (SINGER-ULMER). *Liouvillian solutions and symmetric powers are linked the following way:*

- *The equation $y'' - ry$ has a liouvillian solution if and only if $L^{\circledS 6}$ is reducible.*
- *The equation $L(y) = y''' - a_1 y' - a_0 y = 0$ has a Liouvillian solution if and only if $L^{\circledS 4}$ has order less than 5 or is reducible* AND *(a) $L^{\circledS 2}$ has order 6 and is irreducible* OR *(b) $L^{\circledS 3}$ has a factor of order 4.*

Bibliography

[1] Abramov (S. A.) and Kvashenko (K. Yu.). – Fast algorithms for the search of the rational solutions of linear differential equations with polynomial coefficients. In Watt (Stephen M.) (editor), *Symbolic and algebraic computation*. pp. 267–270. – New York, 1991. Proceedings ISSAC'91, Bonn, Germany.

[2] Bronstein (M.). – On solutions of linear ordinary differential equations in their coefficient field. *Journal of Symbolic Computation*, vol. 13, 1992, pp. 413–439.

[3] Singer (M. F.). – Liouvillian solutions of n-th order homogeneous linear differential equations. *American Journal of Mathematics*, vol. 103, n° 4, 1981, pp. 661–682.

[4] Singer (M. F.). – Liouvillian solutions of linear differential equations with Liouvillian coefficients. *Journal of Symbolic Computation*, vol. 11, 1991, pp. 251–273.

[5] Singer (Michael F.) and Ulmer (Felix). – Galois groups of second and third order linear differential equations. *Journal of Symbolic Computation*, vol. 16, 1993, pp. 1–36.

[6] Singer (Michael F.) and Ulmer (Felix). – Liouvillian and algebraic solutions of second and third order linear differential equations. *Journal of Symbolic Computation*, vol. 16, 1993, pp. 37–73.

[7] Ulmer (Felix). – On Liouvillian solutions of linear differential equations. *Applicable Algebra in Engineering, Communication and Computing*, vol. 2, 1992, pp. 171–193.

Special Polynomials of Ordinary Differential Equations

Jacques-Arthur Weil
GAGE, École Polytechnique

May 30, 1994

[summary by Bruno Salvy]

Abstract

A new algorithm to compute first integrals of quasi-linear ordinary differential equations is presented. Its specialization to the linear case proves useful and in the second order case it leads to an almost rational version of Kovacic's algorithm.

Introduction

Transcendental functions in computer algebra are naturally defined via differential equations. In particular, this makes it possible to test equality between transcendental functions algorithmically, instead of resorting to heuristic rewriting rules. Some functions defined by differential equations occur so often that they have received a special name, and finding an expression in terms of those functions which cancels an equation is called finding a *closed-form* solution. Given an algebraic differential equation of order n or equivalently a polynomial P in $Y, Y', \ldots, Y^{(n)}$, a natural generalization of the idea of closed-form solution is provided by *solutions of order* $r < n$. A function y is a solution of order r of P if there exists a polynomial Q in $Y, Y', \ldots, Y^{(r)}$ such that $Q(y) = 0$ and $P(y) = 0$, while there is no polynomial of order lower than r cancelling y. Finding these solutions is the topic of [9] and [8], of which we give here an overview.

The quest for these solutions is simplified by a well-known *reduction algorithm* [5, p. 5–7] which provides an analogous of Euclidean division in the differential case. If P and Q are two polynomials as above, with $r \leq n$, this reduction proceeds in two steps. First Q is differentiated $n - r$ times, then P is reduced with respect to these derivatives by Euclidean divisions, except that no division by the leading terms are performed. Thus in a finite number of steps, one gets an identity

$$\mathrm{In}(Q)^m \,\mathrm{Sep}(Q)^{n-r} P = \sum_{i=0}^{n-r} \alpha_i Q^{(i)} + R, \tag{1}$$

where R is of order less than r; $\mathrm{Sep}(Q) := \partial Q/\partial Y^{(r)}$ is the *separant* of Q; $\mathrm{In}(Q)$ is the *initial* of Q, i.e., the coefficient of its highest power of $Y^{(r)}$; m is the degree of P in $Y^{(n)}$.

If Q is a polynomial of order $r < n$ cancelling a rth order solution y of P, then the polynomial R in (1) must be 0, since otherwise $R(y) = 0$ with R of order less than r.

At this level of generality, not much is known at present. In the sequel, following [9], we reduce ourselves to solutions of order $n - 1$ of a polynomial P *linear* in $Y^{(n)}$ (Section 1); then to the case of a linear operator P (Section 2); and finally to the case when this linear operator is of order 2 (Section 3).

1. Solutions of order $n-1$

In this section, P is a polynomial of the form

$$P = s(Y, Y', \ldots, Y^{(n-1)})Y^{(n)} + t(Y, Y', \ldots, Y^{(n-1)}).$$

Let Q be any polynomial of order $n-1$. Reducing the polynomial Q' of order n by P leads to

$$\mathrm{In}(P)Q' = \mathrm{Sep}(Q)P + R_1, \tag{2}$$

with R_1 a polynomial of order less than n. This polynomial is then reduced by Q yielding

$$\mathrm{In}^m(Q)sQ' = \mathrm{Sep}(Q)\,\mathrm{In}^m(Q)P + \alpha Q + R, \tag{3}$$

where m is the degree of Q in $Y^{(n-1)}$ and α and R are polynomials of order less than $n-1$.

In view of (3), a polynomial Q of order $n-1$ is called *special* for P if $R = 0$ in (3). In particular, if Q cancels a solution of order $n-1$ of P, it is a special polynomial. Reciprocally, solutions of a special polynomial Q are either solutions of P or solutions of $\mathrm{In}(Q)$ or $\mathrm{Sep}(Q)$.

As a consequence, finding special polynomials of Q is almost equivalent to finding its solutions of order $n-1$. Unfortunately, no general procedure is known to find these polynomials. Even when the degree m of Q is given, finding Q by an undeterminate coefficients method requires some skill. However, when besides m, α in (3) is given, then J.-A. Weil shows in [9] how the method can be reduced to a set of linear differential equations whose rational solutions are the coefficients of Q. Finding these solutions is then a simple routine via Abramov's algorithm [1, 2]. To solve the problem completely, one would need to produce candidates for α algorithmically. Only heuristics or special cases exist at this point.

2. Linear differential equations

In this section P is a polynomial of the form

$$Y^{(n)} + a_{n-2}Y^{(n-2)} + \cdots + a_0 Y.$$

Any linear equation can be reduced into this form, by dividing out by the leading coefficient and then changing the unknown function into $y\exp(-\int a_{n-1}/n)$.

The specificity of the linear equation is that in reduction (2), the remainder R_1 preserves the homogeneous parts of Q. If $Q = \sum Q_i$ is the decomposition of Q as a sum of homogeneous parts,

$$Q' - \mathrm{Sep}(Q)P = \sum Q_i - P\frac{\partial Q_i}{\partial Y^{(n-1)}}$$

is also a decomposition in homogeneous parts. Therefore if Q is a special polynomial, the corresponding α does not depend on Y and it is sufficient to look for monic homogeneous special polynomials.

Given the degree m of a homogeneous special polynomial, write

$$Q = [y^{(n-1)}]^m + f_0[y^{(n-1)}]^{m-1}y^{(n-2)} + \cdots .$$

The leading term of the remainder of Q' mod P is $f_0[y^{(n-1)}]^m$. Thus $\alpha = f_0$, and the system obtained by undeterminate coefficients is linear in all the coefficients except f_0. J.-A. Weil shows in [9] how this leads to a linear system whose solutions with a rational logarithmic derivative is sought. Algorithms to solve this problem efficiently were given by M. Rothstein in 1976, by J. H. Davenport in 1986 and by M. Bronstein in 1990 (see [3] and references therein).

3. Second order and Kovacic's algorithm

By a general Lie-Kolchin theorem (see [6]), a linear differential equation admits Liouvillian solutions if and only if its associated Ricatti equation has an algebraic solution. In the second order case, this Ricatti equation takes the form

$$u' - r + u^2 = 0. \tag{4}$$

Then a special polynomial is one of order 0, i.e., a polynomial whose solutions are algebraic functions solutions of (4). By Kovacic's algorithm [4], the degree of minimal special polynomials can only be one of $\{1, 2, 4, 6, 12\}$. Then the above algorithm applies and the equation it yields for f_0 is the mth symmetric power of the original linear equation. A reason why the second order can be solved efficiently is any special polynomial yields a Liouvillian solution.

More work on reducing as much as possible of the second order case to rational (instead of exponential) solutions of linear differential equations is presented in [8]. As described in the summary of F. Ulmer's talk in these proceedings, there are three cases to consider in order, each step assuming that the previous one failed. At each step, the possible degrees of the minimal special polynomials are known, but their computation may require looking for an exponential solution of $L^{\circledS m}$. The idea in [8] is that in almost every case, one can *choose* m such that a special polynomial is associated with a *rational* solution of $L^{\circledS m}$. Thus, characterizing all possible solutions of the Ricatti equation (e.g., by group-theoretic properties) also characterizes the corresponding special polynomials. The characterization goes as follows.

3.1. Reducible case. If the differential operator can be written $(d/dx - b(x))(d/dx - a(x))y(x)$, then obviously it has an exponential solution (i.e., a solution whose logarithmic derivative is rational). Before looking for such a solution, one should look for rational solutions, which is much easier. Then one can look for a basis of rational solutions of $L^{\circledS 2}$. If it contains only one solution, [8] proves that the special polynomial of degree 2 it induces must factor. If its factors are distinct, they yield two independent solutions of the differential equation. Otherwise, reduction of order leads to the second one. Only if this fails, one should look for exponential solutions.

3.2. Imprimitive case. According to [4], there exists a solution whose logarithmic derivative is algebraic of degree 2. Besides, a careful analysis of the tables of characters of the possible groups is used in [7] to show that the Galois group is imprimitive if and only if $L^{\circledS 4}$ has a rational solution. Such a solution leads to a special polynomial of degree 4. Since at this stage it is impossible that the Ricatti equation has a rational solution, the special polynomial is either irreducible or has factors of degree 2. Consideration of the possible groups shows that either the basis of rational solutions is reduced to one element and the induced special polynomial is the square of the minimal special polynomial or the basis contains two elements and a linear combination of the special polynomials they induce is a perfect square which can be found by resultant and gcd computation.

3.3. Primitive case. This is the case when solutions are algebraic. It is not necessarily a good idea to look for a special polynomial in this case (see the summary of Ulmer's talk). However, rational solutions of $L^{\circledS 6}$, $L^{\circledS 8}$ or $L^{\circledS 12}$ lead to irreducible special polynomial of these degrees.

If none of the above yields a special polynomial, then there is no Liouvillian solution. A byproduct of this method is that algebraic extensions are never necessary except in the reducible case, where examples show that they can be unavoidable. Another useful property of this approach is that it is not limited to equations of the form $Y'' + b(x)Y = 0$, but extends to the more general case, provided the coefficient of Y' is the logarithmic derivative of a rational function (possibly in an

algebraic extension). Then it is not necessary to perform the change of variable suggested at the beginning of Section 2.

Bibliography

[1] Abramov (S. A.). – Rational solutions of linear differential and difference equations with polynomial coefficients. *USSR Computational Mathematics and Mathematical Physics*, vol. 29, n° 11, 1989, pp. 1611–1620. – Translation of the Zhurnal Vychislitel'noi matematiki i matematichesckoi fiziki.

[2] Abramov (S. A.) and Kvashenko (K. Yu.). – Fast algorithms for the search of the rational solutions of linear differential equations with polynomial coefficients. In Watt (Stephen M.) (editor), *Symbolic and algebraic computation*. pp. 267–270. – New York, 1991. Proceedings ISSAC'91, Bonn, Germany.

[3] Bronstein (Manuel). – The transcendental Risch differential equation. *Journal of Symbolic Computation*, n° 9, 1990, pp. 49–60.

[4] Kovacic (Jerald J.). – An algorithm for solving second order linear homogeneous differential equations. *Journal of Symbolic Computation*, vol. 2, 1986, pp. 3–43.

[5] Ritt (Joseph Fels). – *Differential Algebra*. – A.M.S., 1950, *A.M.S. Colloquium*, vol. XXXIII.

[6] Singer (Michael F.). – Liouvillian solutions of n-th order homogeneous linear differential equations. *American Journal of Mathematics*, vol. 103, n° 4, 1981, pp. 661–682.

[7] Singer (Michael F.) and Ulmer (Felix). – Galois groups of second and third order linear differential equations. *Journal of Symbolic Computation*, vol. 16, 1993, pp. 1–36.

[8] Ulmer (Felix) and Weil (Jacques-Arthur). – *Note on Kovacic's algorithm*. – Prépublication n° 94-13, Institut de Recherche Mathématique de Rennes, Université de Rennes 1, France, July 1994.

[9] Weil (Jacques-Arthur). – The use of the special semi-groups for solving differential equations. In *Symbolic and Algebraic Computation*. ACM, pp. 341–347. – New York, 1994. Proceedings ISSAC'94, Oxford, England.

A Universal Constant for the Convergence of the Newton Method

Jean-Claude Yakoubsohn
Université Paul Sabatier, Toulouse

February 28, 1994

[summary by Xavier Gourdon]

Abstract

A new theorem is given concerning the convergence of the Newton method. In this result appears the constant $h_0 = 0.162434\ldots$ which plays a fundamental part in the localization of "good" initial points.

1. Introduction

Given an algebraic equation over $\mathbb{C}$, $P(z) = 0$, it is well-known that the Newton iteration

$$z_0 \in \mathbb{C}, \qquad z_{n+1} = z_n - \frac{P(z_n)}{P'(z_n)} \tag{1}$$

converges to a solution z^* provided the initial value z_0 is sufficiently close to z^*. This iteration is generally used as a refining step in a root finding algorithm to increase the accuracy of the solutions, for example in the exclusion algorithm described in [1]. The problem of giving sufficient conditions on z_0 for (1) to converge is classical. For example, the Newton-Kantorovitch theorem [2, p. 263] states that under the condition

$$2\left|\frac{P(z_0)}{P'(z_0)^2}\right| \cdot \sup_{|z-z_0|<h} |P''(z)| < 1, \tag{2}$$

with some $h > 0$, the Newton iteration is well defined and converges to the unique solution in $|z - z_0| < 2|P(z_0)/P'(z_0)|$ of the equation $P(z) = 0$. This result presents two disadvantages in practice: condition (2) is not expressed at only one point z_0, and the discs of unicity of a solution are generally small. The first result concerning the convergence of Newton method with a punctual criterion is given by Smale in [3]. A new result of this type is given in the following.

THEOREM 1. *Let P be a univariate complex polynomial of degree d. Let $h_0 \simeq 0.162434\ldots$ be the first positive root of the polynomial $4h^3 - 12h^2 + 8h - 1$. Let $z_0 \in \mathbb{C}$ and $h \in [0, h_0]$ such that*

$$\left|\frac{P^{(k)}(z_0)P(z_0)^{k-1}}{P'(z_0)^k}\right| \leq h^{k-1}, \qquad 2 \leq k \leq d. \tag{3}$$

Then (convergence) *the Newton iteration* (1) *converges to a simple solution z^* of the algebraic equation $P(z) = 0$;* (complexity) *the convergence is super-quadratic, that is*

$$|z_{n+1} - z_n| \leq a^n |z_1 - z_0| \left(\frac{h}{a^2}\right)^{2^n - 1},$$

where $a = 2h_0^2 - 4h_0 + 1 \simeq 0.404488\ldots$ *and* $h_0/a^2 \simeq 0.990156\ldots$*;* (set of unicity) *for* $z \in \mathbb{C}$*, define the polynomials in* t

$$L(z,t) = 1 - \sum_{k=1}^{d-1} \frac{|P^{(k)}(z)|}{k!} t^{k-1} \quad \textit{and} \quad \overline{L}(z,t) = tL(z,t) - |P(z)|.$$

Denote by $\ell(z^*)$ *the positive root of* $L(z^*,t)$*. The polynomial* $\overline{L}(z,t)$ *is concave over* $\mathbb{R}$ *and admits either no real roots or two positive roots* $\ell^-(z) \le \ell^+(z)$*. Then each set of form* $|z - z_n| < \ell^+(z_n)$ *for the indices* n *such that* $\ell(z^*) \ge \ell^-(z_n)$ *(this happens for* n *large) contains only one solution of* $P(z) = 0$ *which is* z^**.*

This result generalizes well for algebraic systems [4].

2. Proof of convergence

It is interesting to give a general idea of the proof to understand the origin of the universal constant h_0. Suppose z_0 satisfies conditions (3). A first inequality on $P(z_1)$ is easily derived:

$$|P(z_1)| = \left| P\left(z_0 - \frac{P(z_0)}{P'(z_0)} \right) \right| \le \sum_{k=2}^{d} h^{k-1} |P(z_0)| \le \frac{h}{1-h} |P(z_0)|. \tag{4}$$

Next, we would like z_1 to satisfy conditions (3). Expanding, it is easy to obtain the inequalities

$$\left| \frac{P^{(k)}(z_1) P(z_1)^{k-1}}{P'(z_1)^k} \right| \le h^{k-1} \left(\frac{h}{1-h} \right)^{k-1} \frac{S_{k,d}(h)}{T_d(h)^k}, \qquad 2 \le k \le d$$

where

$$S_{k,d}(h) = \sum_{i=0}^{d-k} \binom{k+i}{i} h^i \quad \text{and} \quad T_d(h) = 1 - \sum_{i=1}^{d-1} (i+1) h^i.$$

Thus, we need $Y_{k,d}(h) = h^{k-1} S_{k,d}(h) - (1-h)^{k-1} T_d(h)^k$ to be negative. It is technical but feasible to show that the polynomials $Y_{k,d}$ have only one positive root $y_{k,d}$, and that they satisfy $Y_{k,d}(h) < 0$ for $0 \le h \le y_{2,d}$. The sequence $y_{k,d}$ is strictly decreasing and tends to the smallest root $h_0 \simeq 0.162434\ldots$ of the polynomial $4h^3 - 12h^2 + 8h - 1$ (therefore it is possible to replace h_0 by $y_{2,d}$ in the theorem). Now, by induction, inequality (4) leads to $|P(z_n)| \le \left(\frac{h}{1-h}\right)^n |P(z_0)|$, showing that $P(z_n) \to 0$ and by continuity, (z_n) converges to a solution z^* of $P(z) = 0$.

3. Conclusion

This result gives a good refining algorithm that fits well with the exclusion method [1]. The result of stability in the theorem also provides good bounds for a classical homotopy method: starting from the roots of a polynomial $Q(z)$, we find the roots of $P(z)$ by finding those of the polynomials $H_t(z) = tP(z) + (1-t)Q(z)$ for successive values of t between 0 and 1.

Bibliography

[1] Dedieu (Jean-Pierre) and Yakoubsohn (Jean-Claude). – Computing the real roots of a polynomial. *Numerical Algorithms*, vol. 4, 1993, pp. 1–24.

[2] Demidovitch (B.) and Maron (I.). – *Éléments de calcul numérique*. – MIR, Moscou, 1979.

[3] Smale (S.). – The fundamental theorem of algebra and complexity theory. *Bulletin of the American Mathematical Society*, vol. 13, 1981, pp. 1–36.

[4] Yakoubsohn (J.-C.). – Approximating the zeros of analytic functions by the exclusion algorithm. *Numerical Algorithms*, vol. 6, n° 1-2, January 1994, pp. 63–88.

Algorithms With Exact Divisions Made Faster

Arnold Schönhage
Institut für Informatik II der Universität Bonn

May 24, 1994

[summary by François Morain]

1. Preamble

And God said:

RULE 1: *Do care about the size of @ !*
RULE 2: *Do not waste a factor of two!*
RULE 3: *Do trust the truth!*
RULE 4: *Do not raise the overall cost for speeding up rare cases – avoid lotteries!*
RULE 5: *Correctness implies termination in due time!*
RULE 6: *Don't forget the algorithms in object designs!*
RULE 7: *Clean results by approximate methods is sometimes much faster!*

and he added: *life will not be easy, since*

IRON RULE: *The development of fast algorithms is slow!*

There are several algorithms that can benefit from exact division in a ring. A typical case is Bareiss' algorithm for performing Gaussian elimination on matrices with integral entries. Another example is the computation of resultants using subresultants. All this and asymptotically fast algorithms are given in [3].

We will describe algorithms for performing exact division of elements in a ring. The setting of these algorithms is given in the recent Bible [2] (from which the Eight Commandments were taken), which contains many fast algorithms for performing arithmetic in various rings or fields. In particular, there is defined the @ symbol, which is pronounced "bounded" and is a synonym for $O(1)$.

2. Exact division

2.1. Theory. The problem is easy to state. Let f and g be two elements of a ring and suppose there exists q in the ring such that $f = qg$. We want to compute q as fast as possible. For instance, if f and g are integers, f with $2m$ words and g with m words, then classical division requires approximately $\gamma_0 m^2$ operations, where γ_0 is the typical constant measuring the cost of a basic operation in the ring.

Jebelean [1] has devised an algorithm for this that requires $\frac{1}{2}\gamma_0 m^2$ operations, by using 2-adic division starting from bottom words. In [2], a similar algorithm is given with same complexity, but starting from the higher words. It was tempting to try to unify the two approaches.

Suppose for instance that f and g are polynomials in $F[y]$, where F is some field. The algorithm of Jebelean on the lowest l coefficients has cost $\gamma_0(l+1)(l+2)/2$ and the algorithm on the upper part has cost $\gamma_0 k(k+1)/2$. Since $l = m-k$, the total cost is minimized for $l = \lfloor m/2 \rfloor$ and $k = \lceil m/2 \rceil$, with a total cost of $\frac{1}{4}\gamma_0 m^2$.

This approach can be extended to the case of integer division with suitable modifications (including a little overlap) and to $\mathbb{Z}[y_1, y_2, \ldots, y_v]$ by interpolation techniques [3].

2.2. Algorithms and examples. Write:

$$f(X) = f_{m-1}X^{m-1} + f_{m-2}X^{m-2} + \cdots + f_0,$$
$$g(X) = g_{n-1}X^{n-1} + g_{n-2}X^{n-2} + \cdots + g_0$$

so that the quotient of f by g is:

$$q(X) = q_{m-n}X^{m-n} + q_{m-n-1}X^{m-n-1} + \cdots + q_0.$$

A Maple implementation of Schönhage's algorithm for finding the coefficients q_{m-n}, $q_{m-n-1}, \ldots,$ $q_{m-n-k+1}$ is

```
# f = f[m-1]X^(m-1) + ... + f[0]
# g = g[n-1]X^(n-1) + ... + g[0]
# q = q[m-n]X^(m-n) + ... + q[0]
# 1 <= k <= m-n+1.
schtabk := proc(fx, gx, x, k)

    local q, i, j, cg, f, g, m, n;

        f:=Pol2Tab(fx, x);
        g:=Pol2Tab(gx, x);
        m:=degree(fx, x)+1;
        n:=degree(gx, x)+1;
        cg:=1/g[n-1];
        for i from m-n by -1 to m-n-k+1 do
            q[i]:=f[i+n-1]*cg;
            for j from n-2 by -1 to m-k-i do
                f[i+j]:=f[i+j]-q[i]*g[j];
            od;
        od;
        sum(q['i']*x^'i', 'i'=m-n-k+1..m-n);
end:
```

This algorithm is the ordinary algorithm for computing the quotient of two polynomials, except that we need to use the coefficients of each intermediary result `f[]` up to degree `m-n-k+1`, which saves some time. The cost of this is easily seen to be $\gamma_0 k(k+1)/2$. Note that this algorithm does not suppose that $g \mid f$, as long as k is not too large.

Jebelean's algorithm, taken from [1] is given below:

```
jebtabl := proc(fx, gx, x, l)

    local q, i, j, cg, f, g, m, n, K;

        f:=Pol2Tab(fx, x);
```

```
        g:=Pol2Tab(gx, x);
        m:=degree(fx, x)+1;
        n:=degree(gx, x)+1;
        cg:=1/g[0];
        K:=m−n+1;
        for i from 0 to l do
            q[i]:=cg*f[i];
            for j from 1 to l−i do
                f[i+j]:=f[i+j]−q[i]*g[j];
            od;
        od;
        sum(q['i']*x^'i', 'i'=0..l);
    end:
```

For the explanation of the algorithm, we refer to the original paper [1]. The cost of this procedure is $\gamma_0 l(l+1)/2$. The main procedure is:

```
    # Assumes f has degree 2m − 1, g has degree m − 1 and f = gq with q of degree m.
    ediv := proc(f, g, x)

      local m, l, k, qj, qs;

        m:=iquo(degree(f, x)+1, 2);
        l:=iquo(m, 2);
        k:=m−l;
        qj:=jebtabl(f, g, x, l);
        qs:=schtabk(f, g, x, k);
        qj+qs
    end:
```

the last missing procedure being:

```
    Pol2Tab := proc(fx, x)

      local f, i;

        for i from 0 to degree(fx, x) do
            f[i]:=coeff(fx, x, i);
        od;
        op(f)
    end:
```

Trying this, we see that:

```
    |\^/|     Maple V Release 2 (Ecole Polytechnique)
._|\|   |/|_. Copyright (c) 1981-1993 by the University of Waterloo.
 \  MAPLE  /  All rights reserved. Maple and Maple V are registered
 <____ ____>  trademarks of Waterloo Maple Software.
      |       Type ? for help.
> f:=X^7+3*X^6+6*X^5+10*X^4+10*X^3+9*X^2+7*X+4:
> g:=X^3+2*X^2+3*X+4:
> ediv(f,g,X);
```

```
                2    3    4
1 + X + X  + X  + X
```

Bibliography

[1] Jebelean (T.). – An algorithm for exact division. *Journal of Symbolic Computation*, vol. 15, 1993, pp. 169–180.

[2] Schönhage (A.), Grotefeld (A. F.), and Vetter (E.). – *Fast algorithms – A multitape Turing machine implementation.* – BI-Wissenschaftsverlag, Mannheim, 1994.

[3] Schönhage (A.), Grotefeld (A. F.), and Vetter (E.). – A new approach to resultant computations and other algorithms with exact division. – 1994. Preprint.

Part 3

Asymptotic Analysis

Travel Inside a "Funny" Complex Differential Equation

Philippe Jacquet
INRIA

December 13, 1993

[summary by Joris van der Hoeven]

Abstract

We consider a functional-differential equation of the form $h_z(z,u) = h(puz,u)h(quz,u)$, where $p,q > 0$ with $p+q = 1$, $h(z,1) = e^z$ and $h(0,u) = 1$. This equation arises when studying the number of phrases in the fundamental parsing algorithm due to Lempel and Ziv. More precisely, it is shown that this problem is equivalent to a problem on digital trees, which reduces to determining the asymptotics of the above equation.

1. The Lempel-Ziv parsing algorithm

The Lempel-Ziv parsing algorithm takes a word as input and partitions it into phrases (blocks) of variable size. In this partition, each new phrase consists of an old phrase, as long as possible, together with a new letter. For instance, the string 11001010001000100 is parsed into (1)(10)(0)(101)(00)(01)(000)(100). Replacing each new phrase by a pointer to the old phrase, together with the new letter, yields a universal data compression algorithm [9]. Other applications are tests of randomness, efficient transmission of data, etc [7, 8, 9].

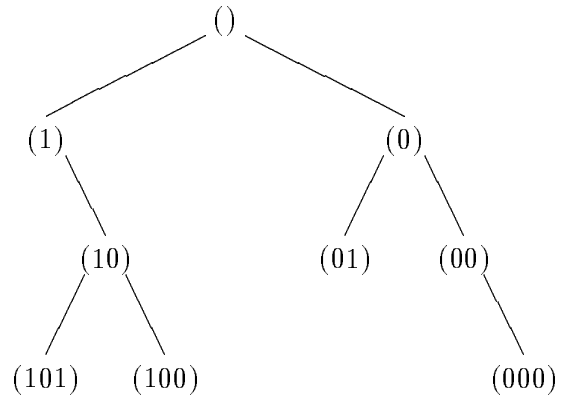

FIGURE 1. The digital tree associated to the string 11001010001000100

Different parameters can be associated to this parsing algorithm, to start with the input size n and the number of phrases $m = M_n$. It is also natural to associate a digital tree to an input word, by interpreting the pointers of the compression algorithm as connections between parents and children. The tree associated to our example string is shown below. The natural size m of a digital

tree is its number of nodes, which corresponds to the number of insertions made and to the number of phrases in the compressing algorithm. Moreover, the internal path length L_m corresponds to the size n of the input word. Therefore, we have

$$M_n = \max\{m | L_m \leq n\}. \tag{1}$$

If we suppose that the insertions are made randomly (this will be made precise), then L_m is a random variable depending on m and from (1) it follows that the random variables M_n in the compression model and L_m in the digital tree model are linked by the formula

$$\Pr(M_n \geq m) = \Pr(L_m \leq n).$$

Furthermore, the relation (1) is known as the renewal equation and from standard probability theory [1], it follows that if L_m has a normal limit distribution, then so has M_n, with $EM_n = n/(EL_n/n)$ and $\operatorname{Var} M_n = \operatorname{Var} L_n/(EL_n/n)^{3/2}$. In the next section we will show that this is actually the case.

The probabilistic model we choose is the asymmetric Bernoulli model. That is, we choose independently letters of a binary alphabet with respective probabilities p and q, where $p, q > 0$ and $p + q = 1$. In the digital tree model, this induces a random insertion of the phrases, which occur during the parsing. Now the inner path length of a digital tree is the sum of the inner path lengths of its two subtrees plus its size. If we note $H_m(u) = E[u^{L_m}]$, this yields

$$H_m(u) = u^m \sum_{k=0}^{m-1} \binom{m-1}{k} p^k q^{m-1-k} H_k(u) H_{m-1-k}(u).$$

Finally, by setting $h(z, u) = \sum_m H_m(u) z^m/m!$, we get our functional equation:

$$h_z(z, u) = h(puz, u)h(quz, u), \tag{2}$$

with $h(z, 1) = e^z$ and $h(0, u) = 1$ as boundary conditions.

2. Analysis of the functional equation

We would like to transform the equation, so that it takes an additive form. If, for a moment, we forget about the derivative, this can be done by studying the logarithm of h and the equation can be rewritten as

$$\ln h(z, u) = \ln h(puz, u) + \ln h(quz, u). \tag{3}$$

We have to verify that the logarithm of h exists. In this simple case this is straightforward and we have the bound $\log h(z, u) = O(z^{k(u)})$, where

$$(pu)^{k(u)} + (qu)^{k(u)} = 1.$$

We remark that $k(u) = 1 - \log u/h + O(\log^2 u)$, where $h = -p \log p - q \log q$ denotes the entropy of the alphabet.

In the original case, it is less straightforward to show that the logarithm exists and to obtain a bound. In fact, we will do this, when u belongs to a real neighbourhood $\mathcal{U}$ of 1 and when z belongs to a polynomial cone $\mathcal{C}(D, \delta) = \{x + iy | x \geq 0 \wedge |y| \leq Dx^\delta\}$, where $D \geq 0$ and $0 \leq \delta < 1$. To put the equation (1) in a form like (3), which is easier to manipulate, we need a new auxiliary function $f(z, u) = h(z, u)/h_z(z, u)$. We obtain

$$f_z(z, u) = 1 - \left(\frac{pu}{f(puz, u)} + \frac{qu}{f(quz, u)} \right) f(z, u) \tag{4}$$

Now f has expected polynomial order $O(z^{1-k(u)})$, which should make this equation easier to study. In fact, we have

THEOREM 1. *There exist a convex polynomial cone $\mathcal{C}(D,\delta)$ of z and a real neighbourhood $\mathcal{U}$ of $u=1$, such that $\log h(z,u)$ exists, and $\log h(z,u) = O(z^{k(u)})$, uniformly for $(z,u) \in \mathcal{C}(D,\delta)\times\mathcal{U}$. Moreover, all derivatives of $\log h(z,u)$ with respect to u exist and are of order $O(z^{k(u)+\varepsilon})$, for any $\varepsilon > 0$.*

PROOF. The theorem is at the heart of the analysis, but its proof is quite involved [6]. We will content ourselves with giving some of the main ideas. Let $\mathcal{D}_m$ be the closed disk of centre 0 and radius ρ^{-m}, where $\rho = \max\{p,q\} < 1$. Then $\mathcal{D}_0 \subseteq \mathcal{D}_1 \subseteq \cdots$ and the disks satisfy the fundamental property

$$z \in \mathcal{D}_{m+1} - \mathcal{D}_m \Rightarrow puz, quz \in \mathcal{D}_m, \tag{5}$$

for $m \geq 0$. Moreover, as u is real, for each convex cone $\mathcal{C}(D,\delta)$, a similar property is satisfied:

$$z \in \mathcal{C}(D,\delta) \Rightarrow puz, quz \in \mathcal{C}(D,\delta).$$

This will make it possible to use induction over the domains $\mathcal{D}_m \cap \mathcal{C}(D,\delta)$. Then by using (4) we obtain an integral formula for $f(z,u)$, when $z \in (\mathcal{D}_{m+1} - \mathcal{D}_m) \cap \mathcal{C}(D,\delta)$, where the integrand contains $f(puz,u)$ and $f(quz,u)$, with $puz, quz \in \mathcal{D}_m \cap \mathcal{C}(D,\delta)$. Estimations of this integral are used to prove the theorem. □

As a consequence of the theorem we can bound the error term in the expansion of $h(z,e^t)$ by using Taylor's formula with integral rest. We obtain

$$h(z,e^t) = \exp\left(z + X(z)t + V(z)\frac{t^2}{2} + O(t^3 z^{k(e^t)})\right).$$

The mean $X(z)$ and variance $V(z)$ can be computed in the Poisson model (using the Poisson generating function $\tilde{h}(z,e^t) = h(z,e^t)e^{-z}$), by using Mellin transform techniques [4]. This yields

$$\begin{cases} X(z) = \frac{z\log z}{h} + O(z), \\ V(z) = \frac{z\log^2 z}{h^2} + Az\log z + O(z), \end{cases}$$

where $A = 0$ in the case $p = q = 1/2$. We therefore obtain normal asymptotics

$$h(z, e^{t/\sqrt{V(z)}})e^{tX(z)/\sqrt{V(z)}} = \exp\left(\frac{t^2}{2} + O(z^{k(u)-3/2})\right).$$

To complete the analysis, it is necessary to translate these asymptotic expansions back to obtain the limit distribution for L_m. This is done by Cauchy's formula. We have

$$E[u^{L_m}] = \frac{m!}{2i\pi}\oint \frac{h(z,u)dz}{z^{m+1}}.$$

Here the contour is a big circle. It is possible to show [5] that the contribution of the part of the contour outside the polynomial cone is exponentially small, and in view of the discussion in Section 1, Jacquet and Szpankowski obtain the following theorem [4, 6]:

THEOREM 2. *In the asymmetric Bernoulli model, M_n has a normal limit distribution of mean $EM_n \sim nh/\log_2 n$ and variance $\operatorname{Var} M_n \sim Ah^3 n/\log^2 n$. In the case of a symmetric Bernoulli model ($p = q = 1/2$), the variance becomes $\operatorname{Var} M_n \sim (C + \phi(\log_2 n))n/\log_2^3 n$, where ϕ is a periodic function of small amplitude and period 1.*

We remark that in the case of a symmetric Bernoulli model, A vanishes, and the following term in the asymptotic expansion gives rise to the oscillating function in the result. We also remark that the theory can be generalized to an alphabet with more letters; this would give functional equations of the type

$$h_z(z,u) = h(p_1 zu, u) \cdots h(p_\nu zu, u),$$

with $\nu \geq 2, p_1, \ldots, p_\nu > 0$ and $p_1 + \cdots + p_\nu = 1$. However, ν may not be equal to 1, because the fundamental property (5) for the domains $\mathcal{D}_m$ fails in this case.

Bibliography

[1] Billingsley (Patrick). – *Probability and Measure*. – John Wiley & Sons, 1986, 2nd edition.

[2] Flajolet (P.) and Sedgewick (R.). – Digital search trees revisited. *SIAM Journal on Computing*, vol. 15, n° 3, August 1986, pp. 748–767.

[3] Flajolet (Philippe) and Richmond (Bruce). – Generalized digital trees and their difference-differential equations. *Random Structures and Algorithms*, vol. 3, n° 3, 1992, pp. 305–320.

[4] Jacquet (P.) and Szpankowski (W.). – A functional equation often arising in the analysis of algorithms on words. – Unpublished conference version.

[5] Jacquet (Philippe) and Régnier (Mireille). – *Normal limiting distributions for the size and the external path length of tries*. – Research report n° 827, Institut National de Recherche en Informatique et en Automatique, April 1988.

[6] Jacquet (Philippe) and Szpankowski (Wojciech). – *Asymptotic behavior of the Lempel-Ziv parsing scheme and digital search trees*. – Research report n° 2212, Institut National de Recherche en Informatique et en Automatique, March 1994.

[7] Lempel (A.) and Ziv (J.). – On the complexity of finite sequences. *IEEE Transactions on Information Theory*, vol. 22, n° 1, 1976, pp. 75–81.

[8] Ziv (J.). – Compression, test of randomness and estimating the statistical model of individual sequences. In Capocelli (R.) (editor), *Sequences*, pp. 366–373. – Springer Verlag, New York, 1990.

[9] Ziv (J.) and Lempel (A.). – A universal algorithm for sequencial data compression. *IEEE Transactions on Information Theory*, vol. 23, n° 3, 1977, pp. 337–343.

Asymptotic Analysis of Finite Differences and Rice Integrals

Philippe Flajolet
INRIA Rocquencourt

February 28, 1994

[summary by Philippe Dumas and Danièle Gardy]

Rice's method is designed to estimate sums

$$Df_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} f_k, \tag{1}$$

where the sequence f_n can be extended as an analytic function $\phi(n)$. The asymptotics of the sequence f_n are assumed to be known, and the problem is to obtain the asymptotic behaviour of the sequence Df_n. The obvious bound

$$|Df_n| \leq 2^n \max_k |f_k|$$

is often disappointing, due to cancellation phenomena and an accurate evaluation of the f_k's cannot provide a direct estimate of the Df_n's. Hence more sophisticated techniques, presented in this talk, are needed. The complete paper is presented in [3].

Many problems in the analysis of algorithms lead to a sequence Df_n. In the sixties, Knuth [4, p. 131] encountered the sum

$$U_n = \sum_{k=2}^{n} \binom{n}{k} \frac{(-1)^k}{2^{k-1}-1}$$

in the study of radix exchange sorting. On can find numerous others examples in the analysis of digital structures [4, p. 501], [1, 6] or conflict resolution in broadcast communications [7].

There are two classical approaches to estimate such alternating sums :

- one can arrange the sum to obtain harmonic sums, which can be tackled by Mellin transforms. This is the standpoint of De Bruijn[1] ;
- Rice proposed a direct approach, which relies on the formula

$$Df_n = \frac{(-1)^n}{2i\pi} \int_{\mathcal{C}} \phi(s) \frac{n!}{s(s-1)\cdots(s-n)}\, ds. \tag{2}$$

The path $\mathcal{C}$ is a contour which encloses the points 0, 1, ..., n, but no singularity of $\phi(s)$. It is assumed that $\phi(s)$ is an analytic function which extends the sequence f_n, and has a polynomial growth at infinity.

[1] based on an original of De Bruijn to Knuth ca. 1965 to be found in the middle pages of Knuth's personal copy of the book *Asymptotic Methods in Analysis*

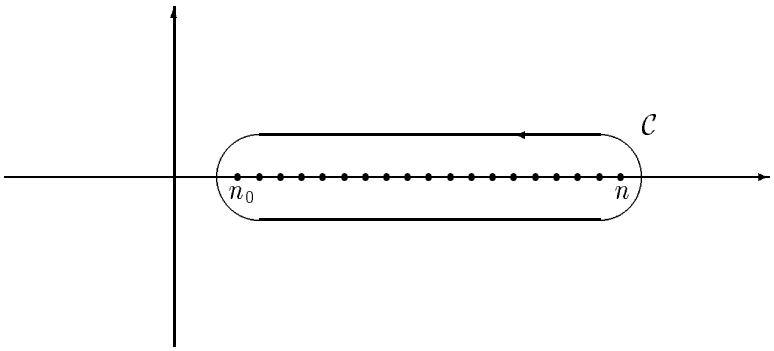

FIGURE 1. A Rice contour.

Here, we develop the second approach, which we call Rice's method. Following Prodinger *et alii*, we write the so-called Rice kernel as follows,

$$[n,s] := \frac{(-1)^n n!}{s(s-1)\cdots(s-n)} = -\frac{\Gamma(-s)\Gamma(n+1)}{\Gamma(n+1-s)}. \tag{3}$$

1. Finite differences

The transformation of sequences (Δ is the forward difference operator, $\Delta f_r = f_{r+1} - f_r$)

$$D : f_n \longmapsto g_n = (-1)^n \Delta^n f_0 = \sum_{k=1}^{n} \binom{n}{k} (-1)^k f_k$$

can be translated into the language of ordinary generating, $F(z) = \sum_n f_n z^n$, or into the language of exponential generating series $f(z) = \sum_n f_n \frac{z^n}{n!}$. In the first case, we obtain essentially Euler's transformation of series

$$G(z) = \frac{1}{1-z} F\left(\frac{-z}{1-z}\right).$$

The second case leads to the formula

$$g(z) = e^z f(-z),$$

which shows the involutive character of the transformation. It is also possible to introduce the Poisson generating series

$$\hat{f}(t) = e^{-t} f(t) = \sum_n f_n e^{-t} \frac{t^n}{n!},$$

which is a simple variant of the exponential generating series. From this point of view, the transformation becomes very simple,

$$g(z) = \hat{f}(-z).$$

Euler transforms and Poisson transforms occur in the analysis of quadtrees [2] and digital structures.

2. Integral representation

The next lemma has been known since the 19th century [5, chap. 8].

LEMMA 1 (RICE'S LEMMA). *Let $\phi(s)$ be an analytic function defined in a neighbourhood Ω of the positive real axis $[0, +\infty)$. Let $\mathcal{C}$ be a contour enclosing the integers $n_0, \ldots, n$, but no singularity of $\phi(s)$. Then*

$$\sum_{k=n_0}^{n} \binom{n}{k} (-1)^k \phi(k) = \frac{1}{2i\pi} \int_{\mathcal{C}} \phi(s) \frac{(-1)^n n!}{s(s-1)\cdots(s-n)}\, ds. \tag{4}$$

The proof is a mere application of Cauchy's residue formula.

The principle of the method is to use hypotheses about the growth of $\phi(s)$ in order to deform the integration contour and obtain an estimate of the sum. More precisely, the function $\phi(s)$ is of polynomial order k if

$$\phi(s) = O(|s|^k) \qquad \text{as } s \to \infty,\ s \in \Omega.$$

With this assumption, the integrand $\phi(s)[n, s]$ tends to 0 when s goes to infinity, if n is large. This permits to modify the path of integration.

3. Rational case

In this section, we consider the basic case of a rational function.

THEOREM 1. *Let $\phi(s)$ be a rational function which is analytic in a neighbourhood of $[n_0, +\infty)$. Then, when n is large enough,*

$$\sum_{k=n_0}^{n} (-1)^k \binom{n}{k} \phi(k) = -\sum_{\omega} \operatorname{Res}\left[\phi(s) \frac{(-1)^n n!}{s(s-1)\cdots(s-n)}, s = \omega\right]. \tag{5}$$

The ω's which appear in the sum are the poles of the integrand not in $[n_0, +\infty)$.

The proof relies on Cauchy's formula used with a contour which is the union of a Rice contour $\mathcal{C}$ and a circle $\mathcal{C}_R$ (cf. Fig. 2). When R tends to infinity, the integral over $\mathcal{C}_R$ tends to 0 provided n is greater than the degree of $\phi(s)$.

Let $s_0 = \sigma_0 + it_0$ be a pole of order r of $\phi(s)$, which we assume not to be a non-negative integer for the sake of simplicity. Then,

$$\operatorname{Res}\left[\phi(s) \frac{(-1)^n n!}{s(s-1)\cdots(s-n)}, s = s_0\right] \underset{n\to\infty}{\asymp} n^{s_0} \log^{r-1} n = n^{\sigma_0} e^{it_0 \log n} \log^{r-1} n.$$

It is possible to be more precise and we refer to [3] for details. The authors obtain an asymptotic equivalent of the type above. Hence Rice's method epitomises a standard asymptotic behaviour mixed with some fluctuations. Moreover it must be pointed out that the rightmost poles of $\phi(s)$ give the most significant part of the asymptotic estimate of Df_n.

EXAMPLE. Let us consider the sum

$$S_n(m) = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^k}{k^m},$$

with m a positive integer. A direct application of the preceding result gives

$$S_n(m) = -\frac{1}{m!} \sum_{k=0}^{m} (-1)^k \binom{m}{k} \Gamma^{(k)}(1) (\log n)^{m-k} + O\left(\frac{\log^m n}{n}\right).$$

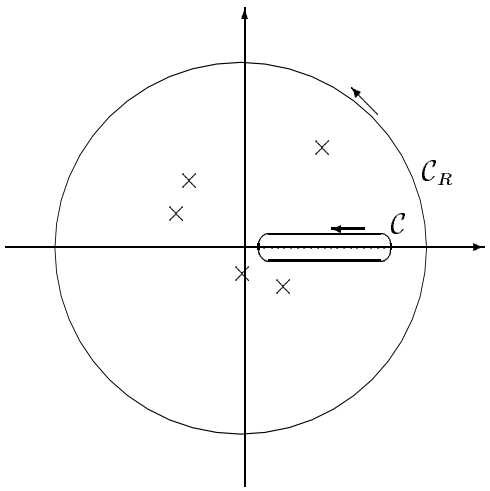

FIGURE 2. For R large enough, all poles (tagged by a cross $\times$) of $\phi(s)$ are inside the circle $\mathcal{C}_R$ but outside the Rice contour $\mathcal{C}$.

EXAMPLE. For the sum

$$T_n = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^k}{k^2+1},$$

the function involved is $\phi(s) = 1/(s^2+1)$ and we expect some term $n^{\pm i}$ to occur. Actually we have

$$T_n = \frac{\pi}{\sqrt{\sinh \pi}} \cos(\log n + \vartheta_0) + o(1).$$

It is to be noticed that T_n remains bounded whereas the central term of the sum is of order $2^n/n^2$.

4. Meromorphic case

The meromorphic case is a mere extension of the rational case.

EXAMPLE. The sum

$$U_n = \sum_{k=2}^{\infty} (-1)^k \binom{n}{k} \frac{1}{2^{k-1}-1}$$

is associated with the meromorphic function

$$\phi(s) = \frac{1}{2^{s-1}-1}, \qquad \text{whose poles are the} \qquad \chi_k = 1 + \frac{2ik\pi}{\log 2}.$$

One integrates over the circles $\mathcal{C}_R$ with center 1 and radius $R = (2k+1)\pi/\log 2$. The circles go between the poles and the function $\phi(s)$ is only of polynomial order on these circles. In this way, one obtains

$$\begin{aligned} U_n &= \frac{n}{\log 2}(H_{n-1} - 1) - \frac{n}{2} + 2 + \frac{1}{2} \sum_{\substack{k \neq 0 \\ k \in \mathbb{Z}}} \frac{\Gamma(n+1)\Gamma(-1+\chi_k)}{\Gamma(n+\chi_k)} \\ &= n \log_2 n + Cn + \frac{n}{\log 2} \sum_{k \neq 0} \Gamma(-\chi_k) e^{2i\pi \log_2 n} + O(\sqrt{n}). \end{aligned}$$

There are many others examples for which we refer to [3] in order to keep this summary short. Let us say simply that Rice's method provides a correspondence between the singularity of $\phi(s)$ and the asymptotic behaviour of Df_n, as summarised below.

Type of singularity		Asymptotic behaviour		
Simple pole	$1/(s-s_0)$	$-\Gamma(-s_0)n^{s_0}$		
Multiple pole	$1/(s-s_0)^r$	$-\Gamma(-s_0)n^{s_0}$	$\dfrac{(\log n)^{r-1}}{(r-1)!}$	
Algebraic singularity	$(s-s_0)^\lambda$	$-\Gamma(-s_0)n^{s_0}$	$\dfrac{(\log n)^{-\lambda-1}}{\Gamma(-\lambda)}$	
Logarithmic singularity	$(s-s_0)^\lambda(\log(s-s_0))^r$	$-\Gamma(-s_0)n^{s_0}$	$\dfrac{(\log n)^{-\lambda-1}}{\Gamma(-\lambda)}$	$(\log\log n)^r$

5. Poisson–Mellin–Newton cycle

A Rice integral may often be written

$$\frac{1}{2i\pi}\int_{c-i\infty}^{c+i\infty}\phi(s)\frac{(-1)^n n!}{s(s-1)\cdots(s-n)}\,ds.$$

For n large, this is approximately

$$\frac{1}{2i\pi}\int_{-c-i\infty}^{-c+i\infty}\phi(-s)\Gamma(-s)n^s\,ds.$$

Hence an idea that we write informally

$$\text{Rice} \approx \text{Mellin}^{-1}.$$

The argument is only heuristic and the right standpoint is the so-called Poisson–Mellin–Newton cycle.

Let us start from a sequence f_n. We associate with it the Poisson series

$$\hat{f}(t) = \sum_{n=0}^{\infty} f_n \frac{e^{-t}t^n}{n!}.$$

The Mellin transform of $\hat{f}(t)$ is

$$\begin{aligned}\hat{f}^*(s) &= \sum_{n=0}^{\infty} f_n \int_0^\infty e^{-t}\frac{t^{s+n-1}}{n!}\,dt\\ &= \sum_{n=0}^{\infty} f_n \frac{\Gamma(s+n)}{n!}\\ &= \Gamma(s)\sum_{n=0}^{\infty} f_n \frac{s(s+1)\cdots(s+n-1)}{n!}.\end{aligned}$$

Hence we have $\hat{f}^*(-s) = \Gamma(-s)\nu(s)$, where $\nu(s)$ is the Newton series

$$\nu(s) = \sum_{n=0}^{\infty}(-1)^n f_n \frac{s(s-1)\cdots(s-n+1)}{n!}.$$

But we can find the coefficient f_n again by the difference operator,

$$f_n = (-1)^n \Delta^n \nu(0).$$

We recognise the expression $D\nu(n)$ and by Rice's lemma we have

$$f_n = \frac{1}{2i\pi} \int_{\mathcal{C}} [n, s]\nu(s)\, ds.$$

Eventually, Rice's transform appears to be the inverse of Mellin transform, composed with Poisson transform.

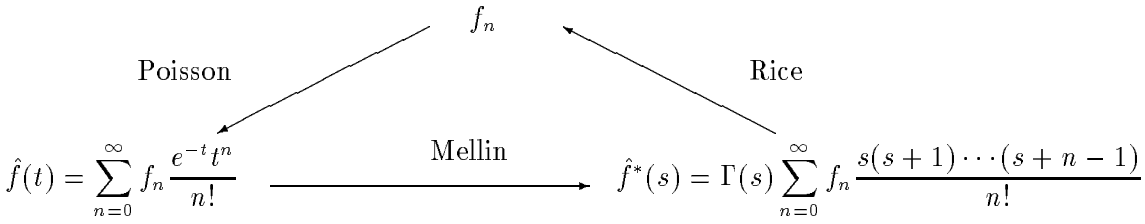

FIGURE 3. The Poisson–Mellin–Newton cycle.

Bibliography

[1] Flajolet (P.) and Sedgewick (R.). – Digital search trees revisited. *SIAM Journal on Computing*, vol. 15, n° 3, August 1986, pp. 748–767.

[2] Flajolet (Ph.), Labelle (G.), Laforest (L.), and Salvy (B.). – *The Cost Structure of Quadtrees.* – Technical Report n° 2249, Institut National de Recherche en Informatique et en Automatique, April 1994.

[3] Flajolet (Philippe) and Sedgewick (Robert). – *Mellin Transforms and Asymptotics: Finite Differences and Rice's Integrals.* – Research Report n° 2231, Institut National de Recherche en Informatique et en Automatique, 1994.

[4] Knuth (Donald E.). – *The Art of Computer Programming.* – Addison-Wesley, 1973, vol. 3: Sorting and Searching.

[5] Nörlund (Niels Erik). – *Vorlesungen über Differenzenrechnung.* – Chelsea Publishing Company, New York, 1954.

[6] Prodinger (Helmut). – How to select a looser. *Discrete Mathematics*, vol. 120, 1993, pp. 149–159.

[7] Prodinger (Helmut) and Szpankowski (Wojciech). – A note on binomial recurrences arising in the analysis of algorithms. *Information Processing Letters*, vol. 46, July 1993, pp. 309–311.

Mellin Transforms and Asymptotics: Harmonic Sums

Xavier Gourdon
INRIA, Rocquencourt

April 25, 1994

[summary by Hsien-Kuei Hwang]

Abstract

This talk gives a general introduction to the asymptotic study of harmonic sums arising in many concrete applications, especially the analysis of algorithms.

1. Introduction

Mellin transform is a precious tool in analytic number theory and in algorithmic analysis as the growth order of the quantity involved is usually polynomial. Given a locally integrable function $f(t)$ on $(0,\infty)$, the Mellin transform $M[f(t);s]$ of f is defined by [11, Ch. III]

$$M[f(x);s] := \int_0^\infty x^{s-1} f(x)\, dx,$$

whenever the integral converges. An essential feature of the function $M[f(x);s]$ is that its domain of analyticity is usually an infinite strip $-\alpha < \Re s < \beta$, the two boundaries $-\alpha, \beta$ being determined, respectively, by the asymptotic behaviours of $f(x)$ when the parameter $x \to 0^+$ and $x \to \infty$. More precisely,

$$\alpha = \sup\{a \,:\, f(x) = O(x^a), x \to 0^+\}, \quad \text{and} \quad \beta = \sup\{b \,:\, f(x) = O(x^{-b}), x \to \infty\}.$$

Thus the inversion formula

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-s} M[f(t);s]\, ds \qquad (-\alpha < c < \beta) \tag{1}$$

offers the flexibility of capturing the asymptotic behaviours of $f(x)$ as x gets small or large by merely shifting the line of integration (here $\Re s = c$) to the left or to the right, respectively, and by collecting the contributions of the singularities encountered (usually poles).

Globally, the important step of "transforming" the given source (function, sequence, etc.) by considering either the associated weighted sums (ordinary/exponential generating function, Dirichlet series, factorial series, etc) or weighted integrals (Laplace, Fourier, Mellin, Hilbert, etc) has the effect of smoothing our "raw data" which becomes more manageable, at least from an analytic point of view.

In view of illustrating the application of Mellin transform to harmonic sums, this talk is mainly example-oriented.

2. Basic Properties

Let $f, M[f;s], \alpha, \beta$ be as in the previous section. Then, cf. [11, Ch. III],

(1) [analyticity] $M[f;s]$ is analytic in the *fundamental strip* $-\alpha < \Re s < \beta$;

(2) [harmonic sums] For any sequences $\{\lambda_k\}$ and $\{\mu_k\}$,

$$M\left[\sum_k \lambda_k f(\mu_k t); s\right] = \left(\sum_k \lambda_k \mu_k^{-s}\right) M[f(x); s];$$

(3) [Riemann-Lebesgue lemma] If $s = \sigma + it$, where $-\alpha < \sigma < \beta$, then

$$\lim_{|t|\to\infty} M[f(x); \sigma + it] = 0;$$

(4) [inversion formula] Under certain regularity conditions,

$$f(x) \sim \pm \sum_{\varpi \in S} \operatorname{Res}\left[M[f(x); s] x^{-s}; s = \varpi\right],$$

as $x \to 0^+$ (the + sign being taken) or $x \to \infty$ (the $-$ sign), where S denotes the set of all singularities (poles) of $M[f(x);s]$ to the left ($x \to 0^+$) or to the right ($x \to \infty$) of the fundamental strip.

3. Two Examples

EXAMPLE. Find the asymptotic behaviour of

$$F(x) = \sum_{k\geq 1} d(k) e^{-kx} \qquad \text{as} \quad x \to 0, \Re x > 0,$$

where $d(k) = \sum_{d|k} 1$ denotes the number of divisors of k. By Property (2),

$$M[F(x); s] = \left(\sum_{k\geq 1} d(k) k^{-s}\right) M[e^{-x}; s] = \zeta^2(s)\Gamma(s) \qquad (\Re s > 1),$$

where ζ denotes Riemann's zeta function and Γ Euler's Gamma function. Standard facts about these two functions [10] and Cauchy's residue theorem lead to the following expansion due to Wigert, cf. [9, p. 163],

$$F(x) \sim \frac{1}{x}\log\frac{1}{x} + \frac{\gamma}{x} + \frac{1}{4} - \sum_{k\geq 1} \frac{B_{2k}^2}{2k\,(2k)!} x^{2k-1} \qquad (x \to 0, \Re x > 0), \tag{2}$$

the B_k being the Bernoulli numbers.

Remark by Hwang. From (2), we obtain

$$\sum_{k\geq 1} d(k) z^k \sim \frac{1}{1-z}\log\frac{1}{1-z} + \frac{\gamma}{1-z} + \cdots \qquad (z \sim 1, |z| < 1),$$

and, in a purely formal way (transferring to coefficient), cf. [4],

$$\sum_{1\leq k\leq n} d(k) = n\log n + (2\gamma - 1)n + \cdots,$$

the determination of the error term constitutes the well-known Dirichlet divisor problem [9, Ch. XII].

EXAMPLE. Find the asymptotic behaviour of

$$F(x) = \sum_{k\ge 0}(1 - e^{-x/2^k}) \qquad \text{as} \quad x \to \infty, |\arg x| < \pi/2,$$

Again, by Property (2), we obtain

$$M[F(x); s] = \left(\sum_{k\ge 0} 2^{ks}\right) M[1 - e^{-x}; s] = -\frac{\Gamma(s)}{1 - 2^s} \qquad (-1 < \Re s < 0).$$

Thus, by the inversion formula (1) and Cauchy's theorem,

$$F(x) = \log_2 x + \frac{1}{2} + \frac{\gamma}{\log 2} + Q(\log_2 x) + R(x) \qquad (x \to \infty, |\arg x| < \pi/2),$$

where $Q(u)$ is a continuous periodic function whose Fourier series is given by

$$Q(u) = -\frac{1}{\log 2} \sum_{k\in\mathbb{Z}\setminus\{0\}} \Gamma\left(\frac{2k\pi i}{\log 2}\right) e^{-2k\pi iu},$$

and $R(x) = O(x^{-M})$ for any $M > 0$.
Remark by Hwang. The error term $R(x)$ can be easily replaced by an asymptotic expansion as follows.

$$R(x) = \frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{\Gamma(s)}{1 - 2^s} x^{-s}\, ds = -\sum_{k\ge 1} \frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \Gamma(s)(2^k x)^{-s}\, ds = -\sum_{k\ge 1} e^{-2^k x},$$

the interchange of the sum and the integral being justified by absolute convergence. This expression for $R(x)$ gives not only an asymptotic expansion but also an exact formula for $F(x)$ as long as $\Re x > 0$.

4. Three Applications to the Analysis of Algorithms

EXAMPLE (AVERAGE HEIGHT OF BINARY (OR PLANAR) TREES). The problem in question [1] (after some reductions) is the asymptotic behaviour of

$$\mu_n := \frac{S_n}{A_n} \qquad \text{as} \quad n \to \infty,$$

where $A_n = \frac{1}{n}\binom{2n-2}{n-1}$ are the Catalan numbers and

$$S_{n+1} = \sum_{k\ge 1} d(k)\left[\binom{2n}{n+1-k} - 2\binom{2n}{n-k} + \binom{2n}{n-1-k}\right].$$

By elementary approximations (using Stirling's formula), we obtain

$$\mu_{n+1} = -2g_0(n) + \frac{4}{n}g_2(n) + O\left(\frac{\log n}{\sqrt{n}}\right),$$

where $g_b(n) = G_b(1/n)$, with $G_b(x) = \sum_{k\ge 1} k^b d(k) e^{-k^2 x}$. Applying Mellin transform to $G_b(x)$ and after some simplification, we obtain [1]

$$\mu_n = \sqrt{\pi n} - \frac{1}{2} + O\left(\frac{\log n}{\sqrt{n}}\right) \qquad (n \to \infty).$$

EXAMPLE (AVERAGE EXTERNAL PATH LENGTH IN A TRIE). Let ℓ_n denote the expected path length of a random trie with n keys. Then [3, 6, 7] ℓ_n satisfies $\ell_0 = \ell_1 = 0$ and

$$\ell_n = n + 2^{-n} \sum_{0 \le k \le n} \binom{n}{k} (\ell_k + \ell_{n-k}) \qquad (n = 2, 3, 4, \ldots).$$

The associated exponential generating function satisfies

$$\ell(z) = \sum_{n \ge 1} \frac{\ell_n}{n!} z^n = z(e^z - 1) + 2\ell(z/2)e^{z/2}.$$

Thus

$$\ell_n = \sum_{2 \le k \le n} \binom{n}{k} \frac{(-1)^k k}{1 - 2^{1-k}} = n \sum_{k \ge 0} \left[1 - (1 - 2^{-k})^{n-1}\right] = n \sum_{k \ge 0} (1 - e^{-n/2^k}) + O(1),$$

and the result of Example 2 gives [6]

$$\ell_n = n \log_2 n + \left(\frac{1}{2} + \frac{\gamma}{\log 2}\right) n + Q(\log_2 n) n + O(1) \qquad (n \to \infty).$$

Two different approaches leading to full asymptotic expansions for ℓ_n are outlined in [2].

EXAMPLE (AVERAGE NUMBER OF CARRY PROPAGATIONS). Let $t(x, y)$ denote the number of carry propagations when adding two numbers x and y and

$$P_{n,k} = 4^{-n} \#\{(x, y) : 0 \le x, y < 2^n, t(x, y) > k\} \qquad (n, k \ge 0).$$

The quantity of interest is $t_n = \sum_{k \ge 0} P_{n,k}$, as $n \to \infty$, which reduces to [5]

$$t_n = \sum_{k \ge 0} [z^n] \left(\frac{1}{1 - z} - \frac{1}{1 - z + z^k/2^{k+1}}\right) = \sum_{k \ge 1} \left(1 - e^{-n/2^k}\right) + O\left(\frac{\log^4 n}{n}\right),$$

where in the last step the localization of the smallest (in modulus) zero of the polynomials $1 - z + z^k/2^{k+1}$ is needed. As in Example 2, we have [5]

$$t_n = \log_2 n + \frac{\gamma}{\log 2} - \frac{1}{2} + Q(\log_2 n) + O\left(\frac{\log^4 n}{n}\right) \qquad (n \to \infty).$$

Remark by Hwang. More calculations show that the O-term can be replaced by an asymptotic expansion of the form

$$\sum_{k \ge 1} n^{-k} \sum_{0 \le j \le k} \pi_{k,j}(\log_2 n) \log^j n,$$

the $\pi_{k,j}(u)$ being periodic functions in u.

5. Some Harmonic Sums

Let $\{\lambda_k\}$ and $\{\mu_k\}$ be two given sequences. Let $f(x)$ be exponentially small at infinity and

$$f(x) \sim \sum_{k\geq 0} f_k x^{\alpha_k} \qquad (x \to 0^+, 0 = \alpha_0 < \alpha_1 < \alpha_2 < \cdots, \alpha_k \to \infty).$$

This assumption implies [11, p. 153] that the function $M[f(x); s]$ admits meromorphic continuation into the whole s-plane with simple poles at $-\alpha_k$, the residues being f_k, for $k = 0, 1, 2, \ldots$. From this fact, we can deduce the following asymptotics as $x \to 0^+$, cf. [2].

(1) [Euler-Maclaurin-Barnes formula]

$$\sum_{n\geq 1} f(nx) \sim \frac{1}{x}\int_0^\infty f(t)\,dt + \frac{f(0)}{2} + \sum_{k\geq 1} f_k \zeta(-\alpha_k) x^{\alpha_k} \qquad (x \to 0^+);$$

(2)

$$\sum_{n\geq 1} (-1)^{n-1} f(nx) \sim \sum_{k\geq 0} (1 - 2^{1+\alpha_k}) f_k \zeta(-\alpha_k) x^{\alpha_k} \qquad (x \to 0^+);$$

(3)

$$\sum_{n\geq 1} f(2^n x) \sim f(0)\log_2 \frac{1}{x} + \frac{f(0)}{2} + \frac{\gamma_f}{\log 2} + \delta(\log_2 x) + \sum_{k\geq 1} \frac{f_k}{1 - 2^{\alpha_k}} x^{\alpha_k} \qquad (x \to 0^+),$$

where

$$\begin{aligned} \gamma_f &= \int_0^1 \frac{f(t) - f(0)}{t}\,dt + \int_1^\infty \frac{f(t)}{t}\,dt \\ \delta(u) &= \frac{1}{\log 2} \sum_{k\in\mathbb{Z}\setminus\{0\}} M[f; 2k\pi i/\log 2] e^{-2k\pi i u}, \end{aligned}$$

the latter being a continuous periodic function of period 1.

Many other types of harmonic power sums can be found in [8, Ch. III].

6. Some Amusing Sums

Let $F(x) = \sum_{m,n\geq 1} f((m^2 + n^2)x)$. To derive an asymptotic expansion for $F(x)$, as $x \to 0^+$, we observe that

$$M[F(x); s] = \left(\sum_{m,n\geq 1} (m^2 + n^2)^{-s}\right) M[f(x); s] = \frac{M[\Theta^2(x); s]}{\Gamma(s)} M[f(x); s].$$

where $\Theta(x) = \sum_{n\geq 2} e^{-n^2 x}$. The singularities of $M[\Theta^2(x); s]$ to the right of the vertical line $\Re s = 1$ (included) are determined by the asymptotic behaviour of $\Theta(x)$ as $x \to 0^+$, for which we apply once again Mellin transform. We thus obtain the functional equation, cf. [9, §2.6],

$$\Theta(x) = \frac{1}{2}\sqrt{\frac{\pi}{x}} - \frac{1}{2} + \sqrt{\frac{\pi}{x}}\,\Theta\left(\frac{1}{x}\right) \qquad (x > 0), \tag{3}$$

the last term being exponentially small as $x \to 0^+$. Once the singularities of the function $M[\Theta^2(x); s]$ are explicited, the asymptotic behaviour of $F(x)$ can easily be derived.

In the same manner, we can consider harmonic sums of the types

$$\sum_{n\geq 1} \lfloor \sqrt{n} \rfloor f(nx), \quad \sum_{n\geq 1} \lfloor \log_2 n \rfloor f(nx), \ldots$$

7. Conclusion

Besides harmonic sums, Mellin transform finds applications to the asymptotics of integrals (especially of convolution type), to the asymptotic behaviour of generating functions, Laplace transform, etc, and to many interesting identities and functional equations like (3).

Bibliography

[1] De Bruijn (N. G.), Knuth (D. E.), and Rice (S. O.). – The average height of planted plane trees. In Read (R. C.) (editor), *Graph Theory and Computing*. pp. 15–22. – Academic Press, 1972.

[2] Flajolet (P.), Régnier (M.), and Sedgewick (R.). – Some uses of the Mellin integral transform in the analysis of algorithms. In Apostolico (A.) and Galil (Z.) (editors), *Combinatorial Algorithms on Words. NATO Advance Science Institute Series. Series F: Computer and Systems Sciences*, vol. 12, pp. 241–254. – Springer Verlag, 1985.

[3] Flajolet (P.) and Sedgewick (R.). – Digital search trees revisited. *SIAM Journal on Computing*, vol. 15, n° 3, August 1986, pp. 748–767.

[4] Flajolet (Philippe) and Odlyzko (Andrew M.). – Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, vol. 3, n° 2, 1990, pp. 216–240.

[5] Knuth (D. E.). – The average time for carry propagation. *Indagationes Mathematicae*, vol. 40, 1978, pp. 238–242.

[6] Knuth (Donald E.). – *The Art of Computer Programming*. – Addison-Wesley, Reading, Mass., 1973, vol. 3: Sorting and Searching.

[7] Mahmoud (Hosam). – *Evolution of Random Search Trees*. – John Wiley & Sons, New York, 1992.

[8] Régnier (M.). – *Évaluation des performances du hachage dynamique*. – Thèse de 3e cycle, Université Paris-Sud, 1983.

[9] Titchmarsh (E. C.). – *The Theory of the Riemann Zeta-function*. – Oxford Science Publications, 1986, second edition. Originally published in 1951, revised by D. R. Heath-Brown.

[10] Whittaker (E. T.) and Watson (G. N.). – *A Course of Modern Analysis*. – Cambridge University Press, 1927, fourth edition. Reprinted 1973.

[11] Wong (Roderick). – *Asymptotic Approximations of Integrals*. – Academic Press, 1989.

Introduction à l'itération des fonctions rationnelles

Jacques Carette
INRIA and Waterloo Maple Software

Lundi 13 décembre

[résumé par Michèle Loday-Richaud]

On s'intéresse aux orbites $\mathcal{O}_f(z)$ des points z de $\mathbb{C}$ sous l'action d'un polynôme ou d'une fraction rationnelle f. Par définition, $\mathcal{O}_f(z)$ est la suite des itérés de z sous f :

$$z_0 = z, \quad z_1 = f(z_0), \ldots, \quad z_{n+1} = f(z_n), \ldots$$

On notera f^n l'itérée n^{e} de f, la puissance étant relative à la composition.

1. Exemples élémentaires

Le cas non trivial le plus simple est celui de $f_0 : z \mapsto z^2$. À l'extérieur du cercle $|z| = 1$ l'orbite d'un point tend vers l'infini en restant sur une spirale "croissante". Sur le cercle l'application est la multiplication de l'angle polaire par 2. À l'intérieur du cercle l'orbite tend vers 0 en restant sur une spirale asymptote à 0.

Ensuite vient $f_\varepsilon : z \mapsto z^2 + \varepsilon$. Il existe une application holomorphe ϕ définie à l'extérieur du disque unité $\mathbb{D}$ qui conjugue f_ε à f_0. À l'extérieur de la courbe $|\phi(z)| = 1$ les orbites s'échappent à l'infini. À l'intérieur les points convergent vers un point fixe attractif, et il existe une application holomorphe sur $\mathbb{D}$ qui conjugue f_ε à $z \mapsto \lambda z$.

2. Quelques définitions

DEFINITION 1. Un *point fixe* est un point z tel que $f(z) = z$.

Un *point périodique de période* n est un point fixe de f^n. À tout point périodique de période n on associe son *multiplicateur*

$$\lambda = (f^n)'(z) = \prod_{i=1}^{n} f'(z_i).$$

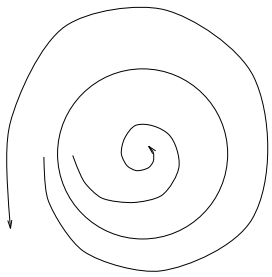

FIGURE 1. Dynamique de $z \mapsto z^2$

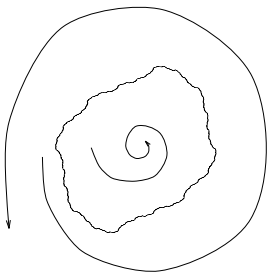

FIGURE 2. Dynamique de $z \mapsto z^2 + \varepsilon$

$$\text{On dit en outre qu'il est } \begin{cases} \text{attractif} & \text{si } |\lambda| < 1, \\ \text{superattractif} & \text{si } \lambda = 0, \\ \text{indifférent} & \text{si } |\lambda| = 1, \\ \text{répulsif} & \text{si } |\lambda| > 1. \end{cases}$$

Un *point critique* est un point z tel que $f'(z) = 0$.

DEFINITION 2. *Ensemble de Julia J_f.* C'est le complémentaire dans $\overline{\mathbb{C}}$ de l'ensemble de Fatou F_f lui-même défini comme l'ensemble des z possédant un voisinage U sur lequel la famille des itérées $f^i : U \to \mathbb{C}$, $i \in \mathbb{N}$ est normale.

Rappelons qu'une famille de fonctions $f_k : U \to \mathbb{C}$ est dite normale si de toute suite on peut extraire une sous-suite qui converge uniformément sur tout compact.

Une approche intuitive de l'ensemble de Julia est donnée par les résultats suivants.

THÉORÈME 1. *J_f est l'adhérence de l'ensemble des orbites périodiques répulsives de f.*

PROPOSITION 1. *Les ensembles de Julia sont non vides, compacts et parfaits (i.e. tout point est un point d'accumulation). Ils sont en général de dimension de Hausdorff non entière.*

PROPOSITION 2. *Soit $z \in J_f$ et U un voisinage de z alors il existe $n \in \mathbb{N}$ tel que $f^n(U) \supset J_f$.*

3. Dynamique sur l'ensemble de Fatou

Limitons-nous à l'étude au voisinage des points périodiques particuliers que sont les points fixes. En effet, d'une part tout point périodique de période n de f est un point fixe de f^n, d'autre part $J_f = J_{f^n}$ pour tout entier n car $f(J_f) = J_f = f^{-1}(J_f)$.

3.1. Cas où z est un point fixe attractif, non superattractif ($|\lambda| < 1$). On peut linéariser f et la conjugante ϕ est analytique ; de plus elle s'étend au bassin d'attraction U de z

$$\begin{array}{ccc} U & \xrightarrow{f} & U \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ \mathbb{C} & \xrightarrow[z \mapsto \lambda z]{} & \mathbb{C} \end{array}$$

3.2. Cas où z est un point fixe répulsif ($|\lambda| > 1$). On peut encore linéariser f.

3.3. Cas où z est un point fixe superattractif ($\lambda = 0$). Alors f est conjugué à une fonction puissance

$$\begin{array}{ccc} U & \longrightarrow & U \\ \downarrow & & \downarrow \\ \mathbb{C} & \xrightarrow[z \mapsto z^n]{} & \mathbb{C} \end{array}$$

où n est le degré topologique de f, c'est-à-dire la somme des degrés de son numérateur et de son dénominateur.

EXEMPLE. $f(z) = z^2 + 1/4$. L'ensemble de Julia J_f dit "le chou-fleur" est une courbe de Jordan. Le point fixe $x_0 = 1/2$ a pour multiplicateur $\lambda = 1$.

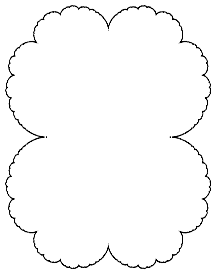

FIGURE 3. L'ensemble de Julia de $z^2 + 1/4$

En conjuguant f par la transformation homographique $z \mapsto Z = 1/(z - 1/2)$ on obtient $F(Z) = Z - 1 + 1/(Z + 1)$. En dehors d'un disque de rayon assez grand, en l'occurrence plus grand que $2\sqrt{2}$ F est injective et voisine de la translation $Z \mapsto Z - 1$.

On peut choisir pour domaine fondamental de F la "bande" comprise entre une verticale γ qui ne coupe pas le disque de rayon $2\sqrt{2}$ et l'image $F(\gamma)$ de celle-ci. On appelle *cylindre d'Ecalle* le cylindre obtenu en recollant γ et $F(\gamma)$.

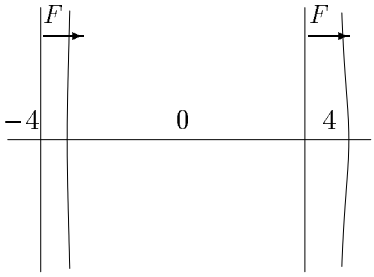

FIGURE 4. Plan des Z

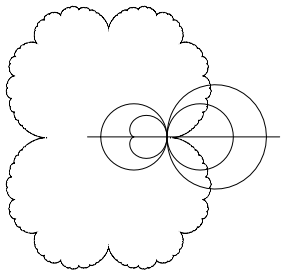

FIGURE 5. Plan des z

La dynamique s'effectue en envoyant un domaine fondamental sur le domaine fondamental contigu. De plus, il existe un domaine, réunion d'un cylindre d'Ecalle et de ses images itérées par f, appelé *pétale de Fatou* sur lequel f est conjugué à la translation $z \mapsto z - 1$

$$\begin{array}{ccc} U & \xrightarrow{f} & U \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle\phi} \\ \mathbb{C} & \xrightarrow[z\mapsto z-1]{} & \mathbb{C} \end{array}$$

Si le multiplicateur $\lambda = e^{2i\pi p/q}$ est une racine de l'unité, le pétale de Fatou est à remplacer par une fleur de Fatou à q pétales, la dynamique envoyant un pétale sur le pétale "suivant".

Le cas où $\lambda = e^{i\pi\theta}$ avec θ irrationnel se sépare en deux suivant les propriétés des dénominateurs des réduites p_n/q_n du développement en fraction continue de θ.

Dans le cas de Siegel, il existe alors des courbes conjuguées à un cercle et l'application f est conjuguée à $z \mapsto \lambda z$ sur leur intérieur appelé *disque de Siegel.*

THÉORÈME 2 (BRUNO-YOCCOZ). *Le point fixe z admet un disque de Siegel si et seulement si*

$$\sum \frac{\log q_{n+1}}{q_n} < +\infty.$$

Le cas complémentaire où $\sum \log q_{n+1}/q_n = +\infty$ est le cas de Cremer.

Les points de Siegel appartiennent à l'ensemble de Fatou, alors que les points de Cremer appartiennent à l'ensemble de Julia.

THÉORÈME 3 (SULLIVAN). *Toute composante connexe U de l'ensemble de Fatou est prépériodique, ce qui signifie qu'il existe k et n entiers tels que $f^{k+\ell n}(U) = f^k(U)$, pour tout $\ell \in \mathbb{N}$.*

La démonstration est très difficile et repose sur la théorie des espaces de Teichmüller.

COROLLAIRE 1. *Si tous les points critiques sont strictement prépériodiques alors $J_f = \overline{\mathbb{C}}$.*

THÉORÈME 4. *Ou bien l'ensemble de Julia J_f est égal à $\overline{\mathbb{C}}$ ou bien il est d'intérieur vide.*

EXEMPLE. $J_f = \overline{\mathbb{C}}$ pour $f = (z^2 - 2)/z^2$ et $\overset{\circ}{J}_f = \varphi$ si f est un polynôme.

4. Cas particulier des polynômes

On a le théorème suivant sur la topologie de J_f.

THÉORÈME 5. *J_f est connexe si et seulement si l'orbite de tout point critique borné est bornée ; J_f est totalement disconnexe si et seulement si l'orbite de tout point critique borné est non bornée.*

Les cas intermédiaires sont horribles.

DEFINITION 3. On considère la famille des polynômes quadratiques $f_c(z) = z^2 + c$. On appelle ensemble de Mandelbrot l'ensemble M des $c \in \mathbb{C}$ pour lesquels J_{f_c} est connexe.

Pour un paramètre c à l'extérieur de M, l'ensemble de Julia J_{f_c} est totalement disconnexe.

THÉORÈME 6 (DOUADY-HUBBARD 1980). *L'ensemble M est compact, connexe et plein (i.e. son complémentaire est connexe).*

THÉORÈME 7 (SHISHIKURA). *La dimension de Hausdorff du bord de M est égale à 2.*

CONJECTURE 1. [1] *L'ensemble M est localement connexe.*

Les avancées de cette conjecture reposent en particulier sur la construction d'un modèle combinatoire pour M (cf. tableaux de Hubbard et Branner en particulier).

[1] Due à Douady et partiellement résolue par Yoccoz.

Part 4

Analysis of Algorithms and Data Structures

Special Limit Distributions for Combinatorial Structures

Michèle Soria
LITP-IBP and INRIA

November 8, 1993

[summary by Joris van der Hoeven]

Abstract

The problem of obtaining asymptotic information about parameters of algorithms can often be reduced to the computation of limit distributions of combinatorial structures. Michèle Soria and Philippe Flajolet have shown that for a large number of combinatorial schemes these limit distributions are normal [1]. However, one also frequently encounters discrete limit distributions. In certain degenerate cases it is even possible to obtain continuous special limit distributions. Here, we are interested in these latter two cases and give some examples. More precisely, we shall study special limit distributions arising from a bivariate generating function of the form $F(uC(z))$.

1. Introduction

Consider the generating function $C(z)$ of some combinatorial structure C. Let $F(C)$ be a new combinatorial structure, obtained by applying a combinatorial construction F to C. Then the information about the number of C-structures "in" an $F(C)$-structure is contained in the bivariate generating function $F(uC(z))$. Usually we take one of the following constructors, which are listed with their associated exponential and ordinary generating functions.

Constructor	Labelled(e.g.f.)	Unlabelled(o.g.f.)
Sequence	$\dfrac{1}{1-uC(z)}$	$\dfrac{1}{1-uC(z)}$
Cycle	$\log\dfrac{1}{1-uC(z)}$	$\displaystyle\sum_{k\geq 1}\frac{\phi(k)}{k}\log\frac{1}{1-u^kC(z^k)}$
Set	$\exp(uC(z))$	$\displaystyle\exp\left(\sum_{k\geq 1}\frac{(-1)^{k+1}}{k}u^kC(z^k)\right)$

We will study

$$\Pr(X_n = k) = \frac{[u^k]F(u)[z^n]C^k(z)}{[z^n]F(C(z))},$$

when n tends to infinity. Here X_n is the random variable giving the number of C-structures in an $F(C)$ structure of size n. We will take k of the form $k = \mu_n + x\sigma_n$, where μ_n and σ_n denote respectively the mean value and the standard deviation of X_n. We will also note $r = \rho_C$ and $R = \rho_F$ the convergence radii of C and F.

Several cases should be distinguished, depending on the sign of $C(r) - R$:

- Sub-critical case ($C(r) < R$), leads to a discrete limit distribution;

- Critical case ($C(r) = R$), leads to a continuous special limit distribution, if $R < \infty$;
- Super-critical case ($C(r) > R$), leads to a normal limit distribution.

In the case of continuous limit distributions, we have different types of theorems, depending on what kind of information we wish to obtain. Classically the following types are distinguished:

- Continuity theorem: $X_n \to X \iff \chi_{X_n}(t) \to \chi_X(t)$;
- Integral limit theorem (GLT): $\Pr(x < (X_n - \mu_n)/\sigma_n < y) \to \int_x^y \omega(t)dt$;
- Local limit theorem (LLT): $\sigma_n \Pr(X_n = \lfloor \mu_n + \lambda\sigma_n \rfloor) \to \omega(\lambda)$;
- Exponential tails: $M_{X_n}(t)$ is uniformly bounded in any interval around 0.

Here $\chi_{X_n} = E(e^{itX_n})$ is the characteristic function of X_n and $M_{X_n} = E(e^{tX_n})$ its moment generating function. As generating functions arising from combinatorial problems are often quite regular, one can usually systematically obtain the four types of theorems by using familiar analytical techniques such as the saddle point method and singularity analysis.

2. The sub-critical case

THEOREM 1. *Suppose that $C(z)$ has an algebraic aperiodic singularity*

$$C(z) = \tau - \gamma(1 - z/r)^\lambda + \cdots,$$

with $0 < \lambda < 1$ and $\tau < R$. Then we have a discrete limit distribution,

$$\Pr(X_n = k) \sim \frac{k f_k \tau^{k-1}}{F'(\tau)} \qquad \text{and} \qquad \mu_n \sim 1 + \frac{\tau F''(\tau)}{F'(\tau)}.$$

The proof runs as follows. The condition $\tau < R$ gives

$$F(C(z)) = F(\tau) - F'(\tau)\gamma(1 - z/r)^\lambda + \cdots.$$

The n-th coefficient $[z^n]F(C(z))$ is obtained by singularity analysis. If k is a constant, $[z^n]C^k(z)$ can also be computed by singularity analysis, when $n \to \infty$. Combining these computations, we obtain

$$\Pr(X_n = k) = \frac{[u^k]F(u)[z^n]C^k(z)}{[z^n]F(C(z))} = \frac{f_k k \tau^{k-1}}{F'(\tau)}.$$

Finally, by the usual formula for μ_n, we have

$$\mu_n = \frac{[z^n]C(z)F'(C(z))}{[z^n]F(C(z))} \sim 1 + \frac{\tau F''(\tau)}{F'(\tau)}.$$

Constructor	$\Pr(X_n = k) \sim$	Law	Example
Sequence	$(1-\tau)^2 k \tau^{k-1}$	Geometric δ	General trees
Cycle	$e^{-\tau}\dfrac{\tau^{k-1}}{(k-1)!}$	Poisson δ	Cayley trees
Set	$(1-\tau)\tau^k$	Geometric	
Partition	$e^{1-\tau-e^\tau}\dfrac{B_k \tau^{k-1}}{(k-1)!}$	Bell δ	
Ordered partition	$(1-\tau)^2 e^{-\frac{\tau}{1-\tau}} k f_k \tau^{k-3}$ $f_k = \sum_p \dfrac{1}{p!}\binom{k-1}{p-1}$	Laguerre δ	

TABLE 1. Application of Theorem 1 for some classical constructors in a labelled environment

λ	α	$\Pr(X_n = xn^\lambda)$	Law	Example
$\frac{1}{2}$	1	$\frac{xe^{-x^2/2}}{\sqrt{n}}$	Raleigh	Random mappings
$\frac{1}{2}$	2	$\sqrt{\frac{2}{\pi n}}x^2e^{-x^2/2}$	Maxwell	Pairs of random mappings
$\frac{1}{4}$	1	$\frac{\Gamma(1/4)}{n^{1/4}}P_{1/4}(x)$	Soria	Extended forests

TABLE 2. Applications of Theorem 2

3. The critical case

THEOREM 2. *Let F be an algebraic-logarithmic function: $F(t) = (1-t)^{-\alpha}\log^\beta[1/(1-t)]$. Suppose that $C(z)$ has an algebraic aperiodic singularity*

$$C(z) = 1 - \gamma(1 - z/r)^\lambda + \cdots, \quad \text{with} \quad 0 < \lambda < 1.$$

Then we have a special continuous limit distribution

$$\Pr(X_n = xn^\lambda) \sim \frac{x^{\alpha-1}\gamma^\alpha}{n^\lambda}\frac{\Gamma(\lambda\alpha)}{\Gamma(\alpha)}P_\lambda(\gamma x), \qquad \text{with} \quad P_\lambda(x) = \sum_{m\geq 0}\frac{(-x)^m}{m!\Gamma(-m\lambda)},$$

where $x = O(1)$. We also have $\mu_n \sim \mu n^\lambda$ and $\sigma_n^2 \sim \sigma^2 n^{2\lambda}$.

PROOF. We have

$$F(C(z)) \sim \frac{\lambda^\beta}{\gamma^\alpha(1 - z/r)^{\lambda\alpha}}\log^\beta\frac{1}{1 - z/r}.$$

The results for μ_n and σ_n are easily obtained by singularity analysis. We must now compute $\Pr(X_n = k)$, for $k = xn^\lambda$, where $x = O(1)$. We have

$$\Pr(X_n = k) = \frac{[u^k]F(u)[z^n]C^k(z)}{[z^n]F(C(z))}$$

and again we use singularity analysis. □

PROPOSITION 1. *$P_\lambda(x)$ is normally convergent for $|x| < x_0$ and hypergeometric, if λ is rational.*

PROOF. If $\lambda = p/q$ is rational, we can write

$$P_{p/q}(x) = \sum_{r=1}^{q-1}P_{p/q}^{(r)}(x), \quad \text{where} \quad P_{p/q}^{(r)}(x) = \sum_{m>0}\frac{(-x)^{mq+r}}{(mq+r)!\Gamma(-pm - rp/q)}.$$

These latter functions are easily seen to be expressible as finite sums of generalized hypergeometric functions. □

As an example, extended forests have the following bivariate generating function:

$$F(uE(z)) = \frac{1}{1 - uE(z)},$$

where $E(z) = 2g(2zg(z))$, with $2g(z) = 1 - \sqrt{1-4z}$. We have $E(z) = 1 - \sqrt[4]{1-4z}$ and

$$P_{1/4}(x) = \frac{x}{4\Gamma(3/4)}{}_0F_2\left(3/4, 1/2\middle|\frac{x^4}{4^4}\right) - \frac{x^2}{4\sqrt{\pi}}{}_0F_2\left(3/4, 5/4\middle|\frac{x^4}{4^4}\right) + \frac{x^3}{8\Gamma(1/4)}{}_0F_2\left(3/2, 5/4\middle|\frac{x^4}{4^4}\right).$$

In her talk, Michèle Soria also considered the super-critical case and the critical case with $R = \infty$. Both cases have normal limit distributions. For more information on the bivariate scheme $F(uC(z))$, see [2].

Bibliography

[1] Flajolet (Philippe) and Soria (Michèle). – Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory, Series A*, vol. 53, 1990, pp. 165–182.

[2] Flajolet (Philippe) and Soria (Michèle). – General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, vol. 114, 1993, pp. 159–180.

Limiting Distributions in Product Schemas

Michèle Soria
LITP-IBP, Université Paris 6

June 6, 1994

[summary by Michèle Soria]

Abstract

We study the limiting distribution of a parameter in product schemas of the type $y(u,x) = g(x)F(uw(x))$, related to classical combinatorial constructions. When $g(x)$ is "negligible", the limiting distribution in $y(u,x)$ is the same as in $F(uw(x))$. On the other hand, when g is of "dominant importance", the limiting distribution is shown to be exclusively discrete, Gaussian or Gamma, according to F and w.

1. Introduction

An important trend in asymptotic combinatorics is to classify limiting distributions appearing in combinatorial schemas according to structural and analytic characteristics of combinatorial constructions. The analysis of functional composition $F(uw(x))$, that translates into generating functions the combinatorial operation of substitution, has been largely investigated [1, 2, 6, 5]. It leads to discrete, or normal, or special distributions, according to analytic properties of F and w.

We shall here consider the case of product schemas $y(x,u) = g(x)F(uw(x))$ studied by Drmota and Soria [3]. Consider for example the two classical results on permutations: the number of cycles of fixed length l is asymptotically Poisson distributed, whereas the limiting distribution of the number of cycles of length $\neq l$ is Gaussian. Let's analyze these results from a "combinatorial schema" standpoint: starting with the construction of permutations as *Sets of Cycles of points* leads to the bivariate schema $y(x,u) = \exp(u \log \frac{1}{1-x})$, hence [5] the Gaussian distribution of the number of cycles (with no restriction) in a random permutation. Marking only cycles of fixed length l gives the product schema $y(x,u) = \exp(\log \frac{1}{1-x} - \frac{x^l}{l} + u\frac{x^l}{l})$, where the factor of dominant importance $\frac{1}{1-x}\exp(-\frac{x^l}{l})$ implies the discrete nature of the distribution, and the *Set* construction in $\exp(u\frac{x^l}{l})$ determines that it is Poisson. On the other hand, the bivariate series for cycles of length $\neq l$ is $y(x,u) = \exp(\frac{x^l}{l}).\exp(u(\log \frac{1}{1-x} - \frac{x^l}{l}))$; here the first factor is negligible, and the limiting distribution is Gaussian as in the unrestricted case.

Given a bivariate series $y(x,u) = \sum y_{nk} x^n u^k$, consider the random variables X_n satisfying $\Pr(X_n = k) = y_{nk} / \sum_k y_{nk}$. We are interested in the asymptotic density of X_n (when $n \to \infty$) in a range around the expected value ($k = \mathrm{E}\, X_n + x\sqrt{\mathrm{Var}\, X_n}$).

For a product schema $y(x,u) = g(x)F(uw(x))$, we thus have to evaluate

$$\frac{y_{nk}}{y_n} = \frac{f_k\,[x^n]g(x)w(x)^k}{[x^n]g(x)F(w(x))}$$

where $f_k = [z^k]F(z)$.

We always assume that the coefficients of the Taylor expansions of $g(x)$, $w(x)$ and $F(w(x))$ can be evaluated, by saddle point method or singularity analysis, and in any case the asymptotic behaviour of the coefficients depends on exact or approximate saddle points.

There are cases where the factor $g(x)$ has no influence on the limiting distribution, i.e. the limiting distribution of $y(x,u)$ is the same as the limiting distribution of $F(uw(x))$. Conversely, $g(x)$ may be of dominant importance, and dictates the limiting distribution of $y(x,u)$ to be either Gaussian or Gamma or discrete. We also investigate some interesting cases where neither $g(x)$ dominates nor is dominated in $y(x,u)$.

The notion of dominance can be formulated in terms of asymptotic behaviour of saddle points. In order to be more precise we introduce the notion of dominance in a product of functions.

DEFINITION 1. Let $f(x)$, $g(x)$ be convergent generating functions with non-negative coefficients. We say that $f(x)$ dominates $g(x)$ if $[x^n]f(x)g(x) \sim g(\zeta_n)[x^n]f(x)$, $(n \to \infty)$, where ζ_n is the saddle point of $f(x)$ defined by $\zeta_n f'(\zeta_n) = nf(\zeta_n)$.

Obviously, if the radius of convergence of the first factor is smaller than that of the second one, the first factor usually dominates. But if both factors have the same radius of convergence, the situation is more involved. However this definition agrees with classical scalings: for example $\exp(\frac{1}{1-x})$ dominates $\frac{1}{(1-x)^\alpha}\log^\beta \frac{1}{1-x}$, and also $\frac{1}{(1-x)^\alpha}$ dominates $\log^\beta \frac{1}{1-x}$, etc.

In a bivariate schema $y(x,u) = g(x)F(uw(x))$, according as the limiting distribution is dictated by the first or second factor, we shall say that g is *dominating* or *dominated* in $y(x,u)$

2. g is dominated in $y(x,u)$

When g is regular at the singular curve of $F(uw(x))$, the limiting distribution is shown to be either Gaussian or discrete. But the situation of simultaneous singularities is more difficult to handle. (In the following we shall refer to *admissible* functions in the sense of Hayman [9], and to *alg-log* functions in the sense of Flajolet and Odlyzko [4].)

THEOREM 1. *Suppose that $F(x)$ is an admissible or alg-log function, with finite radius of convergence R. Let $w(r) = R$ and assume that $w(x)$ and $g(x)$ are regular at $x = r$ (i.e. the radii of convergence of $w(x)$ and $g(x)$ are greater than r). Then $g(x)$ is dominated in $y(x,u) = g(x)F(uw(x))$, and the limiting distribution is Gaussian with mean value $\sim \mu_w(R)^{-1}n$ and variance $\sim \sigma_w^2(R)\mu_w(R)^{-3}n$.*

THEOREM 2. *Suppose that $w(x)$ is an admissible or alg-log function, with finite radius of convergence r such that $w(r) = R$ is finite. Furthermore assume that $F(y)$ is regular at $y = w(r)$ and $g(x)$ is regular at $x = r$. Then $g(x)$ is dominated in $y(x,u) = g(x)F(uw(x))$ and has a discrete limiting distribution, with $Pr[X_n = k] \sim kf_k R^{k-1}/F'(R)$.*

Apart from these two simple cases, the situation where two functions are singular is more complex, and there is probably no general criterion to decide a-priori whether $g(x)$ is dominated or not. Nevertheless we can treat many special cases. For example, if $g(x)$ and $w(x)$ are alg-log functions and $F(w(x))$ has an essential singularity at $x = r$ then $g(x)$ is usually dominated -e.g. if $g(x) = w(x) = \frac{1}{1-x}$ and $F(z) = e^z$, $g(x)$ is dominated in $y(x,u)$)-. Yet, by scaling singularities with the notion of dominance in Definition 1, we can expect the following "general" rule.

RULE 1. Suppose that $g(x)$, $w(x)$, and $F(w(x))$ are admissible or alg-log functions. If $F(w(x))$ dominates $g(x)$, then $g(x)$ is (usually) dominated in $g(x)F(uw(x))$.

3. g is dominating in $y(x,u)$

Alternatively to the previous section, where the factor $g(x)$ has actually no influence on the asymptotic limit distribution of $y(x,u) = g(x)F(uw(x))$, this section is devoted to the case where $g(x)$ is of dominant importance. This means that the saddle point ζ_n of $g(x)$, given by $\zeta_n g'(\zeta_n) = ng(\zeta_n)$, can be used instead of the exact saddle points in the evaluation of the mean, variance and probability. Hence we get, in the range of interest $k \gg\ll \mathrm{E}\, X_n$:

$$\frac{y_{nk}}{y_n} \sim \frac{f_k w(\zeta_n)^k}{F(w(\zeta_n))}.$$

Once again, the case of different radii of convergence is easy to handle.

THEOREM 3. *Suppose that $g(x)$ is admissible or alg-log and has finite radius of convergence r. If $w(x)$ and $F(w(x))$ are regular at $x = r$, then $g(x)$ is dominating in $g(x)F(uw(x))$ and has a discrete limiting distribution with* $\Pr[X_n = k] \sim f_k w(r)^k / F(w(r))$.

In general we can expect a rule of the following kind, which is the converse statement of Rule 1.

RULE 2. Suppose that $g(x)$, $w(x)$, and $F(w(x))$ are admissible or alg-log functions. If $g(x)$ dominates $F(w(x))$ then $g(x)$ is (usually) dominating in $g(x)F(uw(x))$.

One very interesting thing in the dominating case is that there are only few kinds of limiting distributions which can be classified in the following way if $g(x)$ has finite radius of convergence.

THEOREM 4. *Let $g(x)$ be admissible or alg-log, with finite radius of convergence r, and saddle point ζ_n. Suppose that $g(x)$ is dominating in $y(x,u) = g(x)F(uw(x))$, then only four situations can appear:*

a) If $\lim_{x \to r-} w(x) = w(r)$ exists and $F(x)$ is regular at $w(r)$, then X_n has a discrete limiting distribution given by $\Pr[X_n = k] \sim f_k w(r)^k / F(w(r))$.

b) If $\lim_{x \to r-} w(x) = \infty$ and $F(x)$ is entire and admissible, then X_n is asymptotically normally distributed, and $\mathrm{E}\, X_n \sim w(\zeta_n)F'(w(\zeta_n))/F(w(\zeta_n))$.

c) If $\lim_{x \to r-} w(x) = w(r)$ exists and $F(x)$ is admissible and singular at $x = w(r)$, then X_n is asymptotically normally distributed with mean value expressed as in b).

d) If $\lim_{x \to r-} w(x) = w(r)$ exists and if $F(x)$ is an alg-log function (with $\alpha > 0$) which is singular at $x = w(r)$, then X_n is asymptotically Gamma distributed with parameter α and $\mathrm{E}\, X_n \sim \alpha / \log \frac{w(r)}{w(\zeta_n)}$.

4. Combinatorial Schemas

Many typical schemas $y(x,u) = g(x)F(uw(x))$ occurring in combinatorial structures related to the "sequence-of" and "cycle-of" constructions, are discussed in [3]. Beside giving illustrations of sections 2 and 3, we also investigate some cases where g is neither dominated nor dominating in $y(x,u)$.

Consider the product schema $y(x,u) = g(x)\exp(uw(x))$, which underlies the examples given in section 1. In the case of cycles of length l, function g has a finite radius of convergence, and w is an entire function. Thus g is dominating, and the limit distribution is Poisson by Theorem 3-a. On the other hand g is dominated for cycles of length $\neq l$, hence the Gaussian limit law.

Now consider the schema $y(x,u) = g(x)\frac{1}{1-uw(x)}$, and let r be the radius of convergence of $F(w(x)) = \frac{1}{1-w(x)}$. If g is an entire function, it is dominated in $y(x,u)$, whereas if g is exponential with radius of convergence r, it is dominating. An interesting situation (which actually

happens for certain random mapping parameters [8]) arises when g is an alg-log function with radius of convergence r, and $w(r) = 1$: the limit law is shown to be hypergeometric [7, 8].

Bibliography

[1] Bender (Edward A.). – Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory, Series A*, vol. 15, 1973, pp. 91–111.

[2] Canfield (E. Rodney). – Central and local limit theorems for the coefficients of polynomials of binomial type. *Journal of Combinatorial Theory, Series A*, vol. 23, 1977, pp. 275–290.

[3] Drmota (Michael) and Soria (Michèle). – Marking in combinatorial constructions and limit distributions. – Submitted to TCS, November 1993.

[4] Flajolet (Philippe) and Odlyzko (Andrew M.). – Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, vol. 3, n° 2, 1990, pp. 216–240.

[5] Flajolet (Philippe) and Soria (Michèle). – Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory, Series A*, vol. 53, 1990, pp. 165–182.

[6] Flajolet (Philippe) and Soria (Michèle). – General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, vol. 114, 1993, pp. 159–180.

[7] Flajolet (Philippe) and Soria (Michèle). – General combinatorial schemas: Special limit distributions. – Manuscript, January 1994.

[8] Gittenberger (Bernhard). – Local limit theorems for distributions of certain random mapping parameters. – Submitted to Random Structures and Algorithms, January 1994.

[9] Hayman (W. K.). – A generalization of Stirling's formula. *Journal für die reine und angewandte Mathematik*, vol. 196, 1956, pp. 67–95.

Limit Theorems for Combinatorial Structures

Hsien-Kuei Hwang

LIX, École Polytechnique

November 8, 1993

[summary by Michèle Soria]

Abstract

This presentation concerns limiting distributions of parameters—like the number of components—in a variety of combinatorial objects. Under general analytic assumptions on the moment generating function, Hwang obtains complete asymptotic expansions for central and local limit theorems (expressing convergence to a Gaussian law), as well as quantitative estimates for probabilities of large deviations.

1. Introduction

It has been well-known since Gončarov that the number of cycles in a random permutation of size n has a Gaussian limiting distribution, with mean and variance both asymptotic to $\log n$. A similar asymptotic normality result (with a scaling factor of $\log\log n$ instead of $\log n$) was obtained by Erdős and Kac for the number of distinct prime factors of a random integer $\leq n$. These two results belong to different areas and were first proved by different techniques. It is shown here that they are in fact different facets of a common analytic structure.

A recent trend in asymptotic combinatorics is to explain similarity of distributions by similarity of "structure" (see e.g. [1, 3, 6]). In various types of combinatorial schemas, Flajolet and Soria [4, 5] proved a series of central limit theorems of the form

$$\Pr\left\{\frac{\Omega_n-\mu_n}{\sigma_n}<x\right\}\underset{n\to\infty}{\to}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}\,dt.$$

There Ω_n is the number of components in a random object of size n, with mean μ_n and variance σ_n^2. The methods of proof rely on complex analysis for evaluating characteristic functions combined with continuity theorems for establishing convergence to the normal law. Using analytic techniques of probability theory, Hwang [7] gives a precise quantification of asymptotic normality. He obtains full asymptotic expansions for distribution functions and densities (this implies well-quantified convergence rates), together with estimates on probabilities of large deviations from the mean.

2. Central and local limit theorems

The starting point is a general condition for the moment generating function $M_n(s)$ of a sequence $\{\Omega_n\}$ of discrete random variables.

Condition 1. Assume that, uniformly for $|s|\leq\rho$ with $\rho>0$,

$$M_n(s)\equiv\sum_m\Pr(\Omega_n=m)\,e^{ms}=e^{\phi(n)u(s)+v(s)}\left(1+O\left(\frac{1}{\kappa_n}\right)\right),\qquad n\to\infty$$

where $u(s)$ and $v(s)$ are analytic for $|s| \leq \rho$, $u''(0) \neq 0$, and where $\phi(n)$ and κ_n tend to ∞ as $n \to \infty$.

The rate of convergence to the normal distribution results from applying Esseen's Theorem, a standard tool of probability theory (see e.g. [2]), that relates the distance between two distribution functions to the distance between corresponding characteristic functions.

THEOREM 1 (CONVERGENCE RATES). *Under condition 1,*

$$F_n(x) \equiv \Pr\left(\frac{\Omega_n - \mu_n}{\sigma_n} < x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt + O\left(\frac{1}{\kappa_n} + \frac{1}{\sqrt{\phi(n)}}\right),$$

uniformly with respect to x as $n \to +\infty$, where $\mu_n = u'(0)\phi(n)$ and $\sigma_n^2 = u''(0)\phi(n)$.

When convergence is slow (typically in applications the rate is of the order of $n^{-1/2}$, $(\log n)^{-1/2}$ or $(\log\log n)^{-1/2}$), it is useful to have a full asymptotic expansion. In fact, a slightly stronger condition on the function $u(s)$ of Condition 1 ("well-behavedness" of [7]) leads to a precise estimate on the characteristic function around $t = 0$, namely

$$\chi_n(t) = e^{-t^2/2}(1 + \sum P_k(it)/\sigma_n^k).$$

This permits in turn to obtain the asymptotic expansion for densities by means of Fourier inversion.

THEOREM 2 (LOCAL LIMIT THEOREM). *Under Condition 1, and if additionally $u(s)$ is "well-behaved" on $[-i\pi, +i\pi]$,*

$$\sigma_n \ \Pr\left(\frac{\Omega_n - \mu_n}{\sigma_n} = x\right) = \sum_{0 \leq k \leq \nu} \frac{p_k(x)}{\sigma_n^k} + O\left(\frac{1}{\kappa_n} + \frac{1}{\phi(n)^{(\nu+1)/2}}\right),$$

uniformly with respect to x as $n \to +\infty$, where

$$p_k(x) = \frac{d}{dx} P_k(-\Phi)(x) \qquad \text{and} \qquad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt.$$

Obtaining an asymptotic expansion for a central limit theorem is trickier, since the distribution function of Ω_n is a step function as Ω_n is discrete. The jumps at a discrete set of points are reflected by the "saw-tooth" function $\frac{1}{2} - \{x\}$ (with $\{x\}$ denoting the fractional part of x) and its repeated integrals. This leads to an oscillating component in expansions. The proof uses the method of Kubilius [8]. We only quote here a simplified version of the theorem, and refer to [7] for a complete statement.

THEOREM 3 (CENTRAL LIMIT THEOREM). *Under Condition 1, if $u(s)$ is "well-behaved" on the interval $[-i\pi, +i\pi]$,*

$$F_n(x) \sim \Phi(x) + \frac{e^{-x^2/2}}{\sqrt{2\pi}} \sum_{k=1}^{\infty} \frac{\pi_k(x) + \varpi_k(x)}{\sigma_n^k} \qquad (n \to \infty),$$

where $\pi_k(x)$ are polynomials of degree $3k+1$ and $\varpi_k(x)$ are periodic functions.

3. Large deviations

Results of the preceding section deal with the behaviour of Ω_n at a distance $O(\sigma_n)$ from the mean. Probabilities of large deviations predict extreme cases, i.e., the situation when x is allowed to be at a distance $\gg \sigma_n$ from the mean.

It is known that analyticity of a moment generating function around 0 is associated with the occurrence of exponential tails for the corresponding probability distribution by Markov's inequality. Like in [5], the argument may be adapted to the $M_n(s)$, providing $\Pr(|\Omega_n - \mu_n| < x\sigma_n) = O(e^{-cx})$, with $c > 0$, for all $x > 0$. Actually, a much more precise formula can be obtained by using a method of Cramer–Kubilius [8]. It consists of two steps: the integral transform technique of associated distributions, and a saddle-point estimate. The next theorem generalizes Cramer's classical result on large deviations for sums of independent, identically distributed random variables.

THEOREM 4 (LARGE DEVIATIONS FOR CENTRAL LIMIT THEOREM). *Under Condition 1,*

$$\frac{1 - F_n(x)}{1 - \Phi(x)} = e^{\phi(n)Q(x/\sigma_n)} \left(1 + O\left(\frac{x}{\kappa_n} + \frac{x}{\sqrt{\phi(n)}}\right)\right), \qquad x = o(\min\{\kappa_n, \sqrt{\phi(n)}\}),$$

for $x > 0$, where $Q(t)$ is a function analytic at 0 whose coefficients are explicitly computable. A similar estimate holds for the symmetric side of the distribution corresponding to $x < 0$.

Hwang also proves an asymptotic expansion for $\Pr(\Omega_n = m)$, for m lying in the interval $\mu_n \pm o(\sigma_n^2)$. In this case the proof uses the saddle-point method applied to Laplace-type integrals.

THEOREM 5 (LARGE DEVIATIONS FOR LOCAL LIMIT THEOREM). *Let $m = \mu_n + x\sigma_n$ with $x = o(\sqrt{\phi(n)})$. Under the hypotheses of Theorem 4 and with the further assumption that $e^{u(r+it)}/e^{u(r)}$ "behaves like" a characteristic function, one has*

$$\Pr(\Omega_n = m) = \frac{e^{-\frac{x^2}{2} + \phi(n)Q(x/\sigma_n)}}{\sqrt{2\pi\phi(n)u''(0)}} \left(1 + \sum_{1 \le k \le \nu} \frac{P_k(x)}{(\phi(n)u''(0))^{k/2}} + O\left(\frac{x^{\nu+1} + 1}{\phi(n)^{(\nu+1)/2}} + \frac{1}{\kappa_n}\right)\right),$$

where $P_k(x)$ is a polynomial of degree k.

4. Application to combinatorial schemas

Decomposable combinatorial objects are built from *sets*, *sequences* or *cycles* of components. This is known to correspond to functional schemas of the form $P(w, z) = F(w, C(z))$. Here, $C(z)$ is the usual counting generating function of the component objects, and $P(w, z)$ is the bivariate generating function of the composite objects with w marking the number of components.

Let therefore $P(w, z) = \sum_{n,k} p_{nk} w^k z^n$ be such a function, so that p_{nk} is the number of structures of size n with k components; we are concerned with the asymptotic behaviour of the number of components in a random structure of size n whose probability distribution is given by $\Pr(\Omega_n = k) = p_{nk} / \sum_k p_{nk}$. Two major types of schemas are studied.

4.1. Exp-log schemas. These schemas are related to *Set* and *Multiset* constructions and they are already known to lead to Gaussian limit distributions [4]. The general form is $P(w, z) = \exp(wC(z) + S(w, z))$, where $C(z)$ is a function of logarithmic type ($a > 0$ and K a constant)

$$C(z) = a \log \frac{1}{1 - z/\rho} + K + o\left(\frac{1}{\log(1 - z/\rho)}\right) \qquad (z \to \rho, z \notin [\rho, \infty[),$$

and $S(w, z)$ is an analytic function for $|z| < \rho + \epsilon$ and $|w| < 1 + \epsilon'$, for some $\epsilon, \epsilon' > 0$.

By singularity analysis, one gets for the moment generating function

$$M_n(s) = e^{\phi(n)u(s)+v(s)} \left(1 + o\left(\frac{1}{\log n}\right)\right),$$

uniformly for small s when $n \to \infty$, with $u(s) = e^s - 1$, $\phi(n) = a \log n$ and $v(s) = K(e^s - 1) + S(e^s, \rho) - S(1, \rho) + \log(\Gamma(a)/\Gamma(ae^s))$. Hence all the conditions of Theorems 1-5 are fulfilled in this case.

Asymptotic normality of the number of cycles in a permutation or of the number of components in a random mapping provide an illustration of exp-log schemas. Hwang notes that the same process applies to Dirichlet series instead of power series (see Hwang's *"Factorisatio Numerorum"* in this volume). Thus the number of distinct prime factors of a random integer $\leq n$ also fits into this analytic schema.

4.2. Alg-log schemas. Another type of schema from [5],

$$P(w, z) = \frac{1}{(1 - wC(z))^\alpha} \left(\log \frac{1}{1 - wC(z)}\right)^k,$$

is related to *Sequence* and *Cycle* constructions and is again known to lead to Gaussian laws under the conditions: k is a non-negative integer, $\alpha > 0$, and $C(z)$ attains 1 before becoming singular.

By singularity analysis, the moment generating function is shown to have the right form for Theorems 1-5, with

$$u(s) = -\log \frac{\rho(e^s)}{\rho(1)}, \quad \phi(n) = n, \quad v(s) = -\alpha \log \frac{\rho(e^s)C'(\rho(e^s))}{\rho(1)C'(\rho(1))}.$$

Bender's schema for meromorphic functions [1] also fits within this framework.

In conclusion very precise quantitative asymptotic normality results hold for many types of combinatorial objects and number-theoretic functions.

Bibliography

[1] Bender (Edward A.). – Central and local limit theorems applied to asymptotic enumeration. *Journal of Combinatorial Theory, Series A*, vol. 15, 1973, pp. 91–111.

[2] Billingsley (Patrick). – *Probability and Measure*. – John Wiley & Sons, 1986, 2nd edition.

[3] Canfield (E. Rodney). – Central and local limit theorems for the coefficients of polynomials of binomial type. *Journal of Combinatorial Theory, Series A*, vol. 23, 1977, pp. 275–290.

[4] Flajolet (Philippe) and Soria (Michèle). – Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory, Series A*, vol. 53, 1990, pp. 165–182.

[5] Flajolet (Philippe) and Soria (Michèle). – General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, vol. 114, 1993, pp. 159–180.

[6] Gao (Zhicheng) and Richmond (L. Bruce). – Central and local limit theorems applied to asymptotic enumerations IV: Multivariate generating functions. *Journal of Computational and Applied Mathematics*, vol. 41, 1992, pp. 177–186.

[7] Hwang (Hsien-Kuei). – *Théorèmes limites pour les constructions combinatoires et les fonctions arithmétiques*. – Thèse de Doctorat, École Polytechnique, 1994.

[8] Kubilius (J.). – *Probabilistic Methods in the Theory of Numbers*. – American Mathematical Society, Providence, Rhode Island, 1964.

Factorisatio Numerorum, Combinatorial Constructions and Gaussian Limit Distributions

Hsien-Kuei Hwang

LIX, École Polytechnique

April 25, 1994

[summary by Xavier Gourdon and Michèle Soria]

Abstract

This talk gives a general overview of some analytic schemas related to counting the number of components in combinatorial constructions over multiplicative structures.

1. Introduction

Each integer n can be uniquely, up to the order of factors, decomposed into a product of prime factors: $n = \prod_{1\le j\le k} p_j^{a_j}$, with $p_1 < \cdots < p_k$, p_j prime, and $j \ge 1, k \ge 1$. From a combinatorial viewpoint, this decomposition is a multiset of prime numbers (components) whose product equals n. Following this line, the problem of integer factorizations can be extended by considering other sets of components, and other ways of counting the factors. For example two classical problems in "Factorisatio Numerorum" are to count the number of factorizations of n into 2,3,4,..., the order of factors being relevant [5] or irrelevant [7].

Hsien-Kuei Hwang's presentation develops a framework for combinatorial constructions on multiplicative structures and their associated Dirichlet series, which shows a perfect similarity with the now classical results in combinatorics for constructions on non labelled additive structures and their associated power series (cf. [9]).

Enumeration problems and probability distributions of factors then rely on an analytic study of Dirichlet series. Along the same line as for additive combinatorial structures (see e.g. [3, 4]), Hwang analyses statistical properties of the number of factors in analytic schemas of *exp-log* form. Precise results of asymptotic normality are given, using Perron's formula and Hankel contours for evaluating integrals. Some extensions to other classes of analytic schemas are also studied. The content of this talk can be found in [4].

2. Combinatorial Constructions

A multiplicative combinatorial structure $\mathcal{C}$ is a set of objects $\alpha \in \mathcal{C}$ with size $|\alpha|$ such that for all n, the number c_n of objects of size n is finite. The enumerating Dirichlet series $C(s)$ of multiplicative class $\mathcal{C}$ is defined by $C(s) = \sum_{\alpha\in\mathcal{C}} |\alpha|^{-s} = \sum_{n\ge 1} c_n n^{-s}$.

Many combinatorial constructions on multiplicative combinatorial structures, related to integer factorizations, translate into simple forms for the associated Dirichlet series. For example the ordinary factorization on $\mathcal{C}$: $n = |\alpha_1|^{a_1}|\alpha_2|^{a_2}\cdots|\alpha_k|^{a_k}$, $\alpha_i \ne \alpha_j$, $\alpha_j \in \mathcal{C}$, $a_j \ge 1$ is associated to the

Construction	Power series $\sum_{n,k} p_{n,k} w^k z^n$	Dirichlet Series $\sum_{n,k} p_{n,k} w^k n^{-s}$
Multiset	$\exp\Big(\sum_{k\geq 1} \frac{w^k}{k} C(z^k)\Big)$	$\exp\Big[\sum_{k\geq 1} \frac{w^k}{k} C(ks)\Big]$
Set	$\exp\Big(\sum_{k\geq 1} \frac{(-1)^{k-1}}{k} w^k C(z^k)\Big]$	$\exp\Big[\sum_{k\geq 1} \frac{(-1)^{k-1}}{k} w^k C(ks)\Big]$
Sequence	$\frac{1}{1 - wC(z)}$	$\frac{1}{1 - wC(s)}$
Cycle	$\sum_{k\geq 1} \frac{\phi(k)}{k} \log \frac{1}{1 - w^k C(z^k)}$	$\sum_{k\geq 1} \frac{\phi(k)}{k} \log \frac{1}{1 - w^k C(ks)}$

FIGURE 1. Combinatorial constructions and generating functions
For power series $C(z) = \sum_{n\geq 1} c_n z^n$, and for Dirichlet series $C(s) = \sum_{n\geq 2} c_n n^{-s}$. In the bivariate functions, p_{nk} is the number of structures of size n with k components.

multiset construction $\mathcal{M}[\mathcal{C}]$ of $\mathcal{C}$, symbolically defined by $\mathcal{M}[\mathcal{C}] = \prod_{\alpha\in\mathcal{C}}(\epsilon + \alpha + \alpha^2 + \alpha^3 + \cdots)$. This construction translates easily into Dirichlet series, namely

$$M(s) = \prod_{\alpha\in\mathcal{C}} (1 + |\alpha|^{-s} + |\alpha|^{-2s} + |\alpha|^{-3s} + \cdots) = \exp\left(\sum_{k\geq 1} \frac{C(ks)}{k}\right).$$

For example, taking $\mathcal{C} = \mathcal{P}$ the set of prime numbers leads to Euler's equality

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \prod_{p\in\mathcal{P}} (1 + p^{-s} + p^{-2s} + \cdots) = \exp\left(\sum_{k\geq 1} \frac{P(ks)}{k}\right) \tag{1}$$

since every integer can be uniquely factored into a product of prime numbers.

Table 1 presents the Dirichlet series associated with ordinary factorization (Multiset), square-free factorization (Set), ordered factorization (Sequence) and cyclic factorization (Cycle). They show a perfect analogy with the corresponding power series for unlabelled structures. This table actually presents bivariate series $P(w,z) = \sum_{n,k} p_{n,k} w^k z^n$ and $P(w,s) = \sum_{n,k} p_{n,k} w^k n^{-s}$ where variable w, considered additively, marks the number of $\mathcal{C}$-components.

3. Analytic schemas

Multiplicative compositions lead to analytic schemas with different types of singularities, according to the construction and the Dirichlet series of the class of components.

If the components are $2, 3, 4, \ldots$ then $C(s) = \sum_{n\geq 2} n^{-s}$ is equal to $\zeta(s) - 1$, which is known to have a simple pole at $s = 1$. On the other hand, if the components are prime numbers, Möbius inversion applied to Euler's equality (1) gives $P(s) = \sum_{p\in\mathcal{P}} p^{-s} = \sum_{k\geq 1} \frac{\mu(k)}{k} \log \zeta(ks)$, hence a logarithmic singularity at $s = 1$ for $P(s)$. We are thus faced with Dirichlet series having a polar or a logarithmic singularity. For multiset and set constructions, this leads naturally to analytic schemas of the form

$$P(w,s) = e^{wW(s)}\Psi(w,s), \qquad W(s) = \frac{K}{s-\rho} + H(s) \quad \text{or} \quad W(s) = K \log \frac{1}{s-\rho} + H(s), \tag{2}$$

where $\Psi(w,s)$ and $H(s)$ satisfy some regularity conditions, whereas for sequence and cycle constructions, the schemas become

$$(3) \qquad P(w,s) = \frac{K(w)}{s-\rho(w)} + H(w,s) \qquad \text{and} \qquad P(w,s) = \log\frac{1}{s-\rho(w)} + H(w,s).$$

3.1. Enumeration of factorizations. Given $F(s) = \sum f_n n^{-s}$ the Dirichlet series associated with some factorization problem (e.g. in "Factorisatio Numerorum", $F(s) = (2-\zeta(s))^{-1}$ or $F(s) = \exp(\sum(\zeta(ks)-1)/k)$ counts the number of ordered or unordered factorizations of n into integers ≥ 2), the question is to find the asymptotic behaviour of the summatory function $A(x) = \sum_{1\leq n\leq x} f_n$.

When $F(s)$ has a polar singularity at $s=\rho$, with some further conditions, Ikehara showed [5] that $A(x) \sim Kx^{\rho}$, and Delange [1] extended this result for $F(s)$ with logarithmic singularity. These results, whereas with different techniques of proof, can be brought to Singularity Analysis [2] for evaluating the coefficients of a power series with algebraic and logarithmic singularities.

On the other hand, for exponential singularities, as in $F(s) = \exp(\frac{1}{s-1})$, the saddle-point method applies (as well as in the case of power series) to estimate the summatory function $A(x)$ [6, 7].

3.2. Statistical properties of the number of factors. Considering the bivariate schemas $P(w,s)$ arising from multiplicative compositions where factors are marked by w, we want to asymptotically characterize the distribution of the number of factors : mean value, variance and other moments; central and local limit theorems, probabilities of large deviations;...

More formally, from a Dirichlet series $P(w,s) = \sum_{n\geq 1} P_n(w) n^{-s}$, where $P_n(w)$ are polynomial in w with nonnegative coefficients, we shall study statistical properties of the random variable ξ_n counting the average number of objects of size $\leq n$ with a given number of component, precisely defined by

$$\Pr(\xi_n = m) = \frac{\sum_{1\leq k\leq n}[w^m]P_k(w)}{\sum_{1\leq k\leq n} P_k(1)}$$

where $[w^m]P_k(w)$ denotes the coefficient of w^m in $P_k(w)$.

Each analytic schema described before leads to a Gaussian limit distribution [4]. In the following, we concentrate on the exp-log schema.

4. Exp-log class

The exp-log schema corresponds to bivariate generating functions of the form

$$P(w,s) = \sum_{n\geq 1} P_n(w) n^{-s} = e^{wW(s)}\Psi(w,s).$$

Here, $W(s)$ is a Dirichlet series with non negative coefficients with an abscissa of convergence $\rho > 0$ and can be written

$$W(s) = K\log\frac{1}{s-\rho} + H(s), \qquad K > 0$$

where $H(s)$ is analytic in

$$\Delta(\rho, c) = \{s \mid s = \sigma + it, \sigma \geq \rho - c/V(t)\}$$

with $V(t) = \log(|t|+3)$ and $c > 0$. Moreover $\Psi(w,s)$ is holomorphic for $\Re(s) \geq \rho - \delta$ with a $\delta > 0$ and for $|w| \leq \eta$. Some growth conditions on $P(w,s)$ are needed in the domain $\Delta(\rho,c)\setminus[\rho-c,\rho]$. Roughly speaking, the Dirichlet series $s \mapsto P(w,s)$ has its dominant singularity at $s=\rho$, near which it behaves like $\exp(Kw\log\frac{1}{1-s})$.

From a Perron-like formula, these assumptions lead to an asymptotic expansion of the summatory function $A(x,w) = \sum_{1\le k\le x} P_k(w)$ as $x \to +\infty$

$$A(x,w) = x^{\rho}(\log x)^{Kw-1}\left[\sum_{j=0}^{\nu}\frac{\Upsilon_j(w)}{\Gamma(Kw-j)(\log x)^j} + O((\log x)^{-\nu-1})\right]$$

uniformly for $|w| \le \eta$, where $\Upsilon_j(w)$ is the j-th coefficient of the expansion of $e^{wH(s)}\Psi(w,s)/s$ at $s=\rho$. Thanks to Selberg's method [8], this expansion translates to coefficients, giving

$$\sum_{1\le k\le x}[w^m]P_k(w) \approx \frac{x^{\rho}}{\log x}\left(\sum_{j\ge 0}\frac{\kappa_{m,j}(K\log\log x)}{(\log x)^j}\right),$$

where the $\kappa_{m,j}$ are polynomials of degree $m-1$ defined from the $\Upsilon_j(w)$.

From this last expression and thanks to general theorems by Hwang ([4], and see also *Limit Theorems for Combinatorial Structures* in these proceedings), it is possible to derive a full asymptotic expansion of the mean, variance and other moments, to show that the limit distribution is Gaussian and to give its full asymptotic expansion (for the central and local limit theorem), and study large deviations.

Application. Studying the random variable Ω_n counting the number of distinct divisors of a random integer $k \in [1,n]$, that is $\Pr(\Omega_n = m) = \frac{1}{n}|\{k : 1 \le k \le n, \omega(k) = m\}|$ where $\omega(k)$ denotes the number of distinct prime factors of k, leads to a bivariate generating function $P(w,s)$ of exp-log type. The general results of the corresponding schema gives, among others,

$$E(\Omega_n) \sim \log\log n, \quad \mathrm{Var}(\Omega_n) \sim \log\log n$$

and a Gaussian limit distribution as first stated by Erdös-Kac (1940).

Bibliography

[1] Delange (H.). – Généralisation du théorème de Ikehara. *Annales Scientifiques de l'École Normale Supérieure*, vol. 71, 1954, pp. 213–242.

[2] Flajolet (Philippe) and Odlyzko (Andrew M.). – Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, vol. 3, n° 2, 1990, pp. 216–240.

[3] Flajolet (Philippe) and Soria (Michèle). – Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory, Series A*, vol. 53, 1990, pp. 165–182.

[4] Hwang (Hsien-Kuei). – *Théorèmes limites pour les constructions combinatoires et les fonctions arithmétiques.* – Thèse de Doctorat, École Polytechnique, 1994.

[5] Ikehara (S.). – On Kalmar's problem in "factorisatio numerorum". *Journal of the Physico-Mathematical Society of Japan*, vol. 23, 1941, pp. 767–774.

[6] Oppenheim (A.). – On an arithmetic function (II). *Journal of the London Mathematical Society*, vol. 2, 1927, pp. 123–130.

[7] Szekeres (G.) and Turán (P.). – Über das zweite Hauptproblem der "factorisatio numerorum". *Acta Litterarum ac Scientiarum, Szeged*, vol. 6, 1932, pp. 143–154.

[8] Tenenbaum (Gérald). – *Introduction à la Théorie Analytique et Probabiliste des Nombres.* – Institut Élie Cartan, Université de Nancy I, 1990.

[9] Vitter (Jeffrey Scott) and Flajolet (Philippe). – Analysis of algorithms and data structures. In van Leeuwen (J.) (editor), *Handbook of Theoretical Computer Science*, Chapter 9, pp. 431–524. – North Holland, 1990.

Average-Case Analysis of Pattern-Matching

Mireille Régnier
INRIA-Rocquencourt

November 29, 1993

[summary by Jean-Marc Steyaert]

Knuth-Morris-Pratt algorithm for pattern-matching has been long studied as well as its variants. Its worst-case behaviour has been analyzed in several implementations and upper-bounds on the delay have been given. This talk gives precise estimates for the average-case behaviour under a variety of probabilistic models: uniform, Bernoulli, Markov, both for patterns and texts. The average complexity happens to be linear in all cases and the linearity constant K can be precisely computed thus allowing full comparison with other algorithms.

The results are obtained by an algebraic and language theoretic approach. Basically variations around the average-case behaviour are due to overlapping subpatterns. These overlaps are formally described by means of formal language theory and in this particular case of context-free grammars.

Average costs can then be expressed in terms of formal power series that satisfy quasi-algebraic equations: perturbative terms happen to be almost neglectible. With some use of computer algebra it is then possible to determine precisely expectation and variance for a number of strategies.

Bibliography

[1] Régnier (Mireille). – Average performance of Morris-Pratt-like algorithms, 1993.

Random Polynomials and Factorization Algorithms

Xavier Gourdon
INRIA Rocquencourt

October 4, 1993

[summary by Henry Crapo]

Introduction

We study the multiplicative structure of the set $\mathbb{F}_q[x]$ of polynomials with coefficients in a finite field $\mathbb{F}_q$ of q elements. The aim is to develop a unified treatment of probabilistic properties (mean, standard deviation, distribution) of the major parameters relevant to polynomial factoring algorithms. Several of the analyses are already known but put here in a common perspective, while some others appear to be new (smallest degree, largest degree, probability that the distinct degree factorization is complete).

The story for polynomials of large degree over some fixed field $\mathbb{F}_q$ goes as follows: A random polynomial has with high probability about $\log n$ irreducible factors, the distribution of the number of factors being Gaussian in the limit, with exponential tails. The asymptotic probability that it is square–free is between $1/2$ ($q = 2$) and 1 ($q = \infty$) while the degree of its square free part is $n - \mathcal{O}(1)$ on average. In other words, a random polynomial is expected to have only a very few repeated factors of very small degree. The number of factors of a fixed degree r is approximately Poisson distributed with parameter $1/r$. The degrees of the smallest and largest irreducible factors are on average about $0.5614 \log n$ and $0.6243\, n$. The distinct degree factorization completely factors a polynomial of large degree with probability that varies between 39% ($q = 2$) and 56% ($q = \infty$).

1. Basic equations

Let $\mathcal{S}$ be a class (species) of combinatorial structures. A class $\mathcal{P}$ of structures is said to be *decomposable* over $\mathcal{S}$ if every element of $\mathcal{P}$ is uniquely expressible as a multiset of elements of $\mathcal{S}$. For example, the cycle-structure of permutations is a multi-set of cycles, and the unitary polynomials over a finite field $\mathbb{F}_q$ are multisets of irreducible unitary polynomials.

Let $\mathcal{I}$ be a class of basic objects of various integral weights with $|\omega|$ denoting the weight of $\omega \in \mathcal{I}$, then

$$I(z) = \sum_{\omega \in \mathcal{I}} z^{|\omega|} = \sum_n I_n z^n,$$

where I_n is the number of objects in $\mathcal{I}$ having weight n. The corresponding generating functions for finite subsets or multisets of $\mathcal{I}$ are respectively

$$Q(z) = \prod_{\omega \in \mathcal{I}} (1 + z^{|\omega|}) = \prod_{n=1}^{\infty} (1 + z^n)^{I_n}, \qquad P(z) = \prod_{\omega \in \mathcal{I}} (1 - z^{|\omega|})^{-1} = \prod_{n=1}^{\infty} (1 - z^n)^{-I_n}.$$

Polynomials. Take $\mathcal{I}$ to be the collection of all monic irreducible polynomials over a finite field $\mathbb{F}_q$, with weight being degree. Let $P(z), Q(z)$, defined above, be the generating function of all monic polynomials, monic square-free polynomials, respectively. Since $P_n = [z^n]P(z)$ has value q^n, we have $P(z) = (1 - qz)^{-1}$, and the first relation implicitly determines I_n. Taking logarithms and applying Möbius inversion to the resulting expression, we obtain

$$I(z) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \log \frac{1}{1 - z^k}, \qquad I_n = \frac{q^n}{n} + \mathcal{O}(\frac{q^{n/2}}{n}),$$

so that a fraction extremely close to $1/n$ of the polynomials of degree n are irreducible. These relations imply that $I(z)$ has only one dominant singularity at $z = 1/q$, around which $I(z) \sim \log[1/(1 - qz)]$. Thanks to the singular part of $I(z)$, we will be able to derive several asymptotic formulæ by singularity analysis [4].

The Euler relation $1/(1 - z) = (1 + z)(1 + z^2)(1 + z^4)\cdots$ applied to the infinite products for $P(z), Q(z)$ entails $P(z) = Q(z)\cdot Q(z^2)\cdot Q(z^4)\cdots$, an identity that corresponds to grouping repeated factors according to the binary representations of their multiplicities. Thus $Q(z) = P(z)/P(z^2)$ so that

$$Q(z) = \frac{1 - qz^2}{1 - qz} \qquad \text{and} \qquad Q_0 = 1, \quad Q_1 = q, \quad Q_n = q^{n-1}(q - 1)\ (n \geq 2).$$

The permutation model. The joint distribution of degrees in the prime decomposition of a random polynomial over $\mathbb{F}_q$ having degree n admits as a limit, as the cardinality q of the base field tends to infinity (n staying fixed!), the joint distribution of cycle lengths in random permutations of size n. This gives rise to a useful heuristic: probabilistic properties of polynomial factorizations often have a shape resembling that of corresponding properties of the cycle decomposition of permutations.

2. Number of irreducible factors

The number of monic irreducible factors of a monic polynomial in $\mathcal{P} = \mathbb{F}_q[x]$ is an additive parameter with bivariate generating function

$$P(z, u) = \prod_{p \in \mathcal{P}} (1 - uz^n)^{-I_n}.$$

From standard generating function techniques, the total number of monic irreducible factors in all monic polynomials of degree n is the coefficient $[z^n]T(z)$ where

$$T(z) = \frac{\partial}{\partial u} P(z, u)\Big|_{u=1} \quad = \quad P(z) \cdot \left(\sum_{k=1}^{\infty} I(z^k) \right).$$

Using Dirichlet convolution and singularity analysis we find the function $T(z)$ has only one dominant singularity at $z = 1/q$ around which

$$T(z) \sim \frac{1}{1 - qz} \cdot \log \frac{1}{1 - qz} \qquad (z \to \frac{1}{q}).$$

By transfer lemmas [4], the total number of monic irreducible factors in all the polynomial of degree n satisfies $T_n \sim q^n \log n$ as $n \to \infty$, so that the mean number of irreducible factors is asymptotically $\log n$. Pretty much everything is known regarding this parameter. The variance is known to be asymptotically $\log n$, and once normalized, the distribution is Gaussian in the limit [5] with exponential tails [6]. A local limit theorem also holds and results from the general theorem of Gao and Richmond [7].

3. Number of irreducible factors of fixed degree

THEOREM 1. *Let r be a positive integer. Let $\Omega_n(r)$ be the random variable counting the number of irreducible factors of degree r of a random polynomial of degree n over $\mathbb{F}_q$, each factor being counted with its order of multiplicity.*

(1) *The mean value $\mu_n(r)$ and variance $\sigma_n(r)^2$ of $\Omega_n(r)$ are asymptotically, as n tends to infinity*

$$\mu_n(r) \underset{n\to\infty}{\sim} \frac{I_r q^{-r}}{1-q^{-r}} \underset{q\to\infty}{\sim} \frac{1}{r}, \qquad \sigma_n(r)^2 \underset{n\to\infty}{\sim} \frac{I_r q^{-r}}{(1-q^{-r})^2} \underset{q\to\infty}{\sim} \frac{1}{r}.$$

(2) *For any fixed integer k,*

$$\Pr\{\Omega_n(r)=k\} \underset{n\to\infty}{\sim} (1-q^{-r})^{I_r} q^{-kr} \binom{I_r+k-1}{k} \underset{q\to\infty}{\sim} e^{-1/r}\frac{r^{-k}}{k!}.$$

The distribution of $\Omega_n(r)$ is approximately Poisson with parameter $1/r$.

The degree of the non-squarefree part of a polynomial in $\mathbb{F}_q[x]$ has order a small constant, this constant furthermore tends to zero as q goes to infinity. The parts of degree r have a size that decreases roughly geometrically (in q^{1-r}) with r.

4. Extreme degrees of irreducible factors

THEOREM 2. *The highest degree of the irreducible factors of a random polynomial of degree n over $\mathbb{F}_q$ has expectation asymptotic to Cn where C is a constant not depending on q, namely*

$$C = \int_0^\infty \left[1 - \exp\left(-\int_x^\infty \frac{e^{-t}}{t}\,dt\right)\right] dx \approx 0.62432965.$$

This constant C had already surfaced in [15] as the limit of ℓ_n/n, ℓ_n denoting the expected length of the longest cycle in a random permutation of n elements. The result is consistent, since the permutation model is the limit of the polynomial of $\mathbb{F}_q$ model as $q \to \infty$.

Numerical experimentations confirm our result. For $q = 2$ and degree 200 and 400 yield $0.62433383\ldots$ (thanks to Romberg convergence acceleration process), approximating the constant C with an error less than 10^{-5}. The proof looks like the one we find in [4] regarding the longest cycle in permutations.

THEOREM 3. *The probability that all irreducible factors of a random polynomial in $\mathbb{F}_q[x]$ of degree n have degree more than r is asymptotically, as $n \to \infty$,*

$$\prod_{j\le r}\left(1-\frac{1}{q^j}\right)^{I_j} \underset{q\to\infty}{\sim} e^{-H_r} \underset{r\to\infty}{\sim} \frac{e^{-\gamma}}{r}.$$

The expected degree should be approximately $\sum_{r=1}^{n} \frac{e^{-\gamma}}{r} \sim e^{-\gamma}\log n$, in accordance with the random permutation model.

THEOREM 4. *The smallest degree of the irreducible factors of a random polynomial of degree n over $\mathbb{F}_q$ has mean asymptotic to $e^{-\gamma}\log n$ where γ denotes Euler's constant. We have $e^{-\gamma} \approx 0.56145948$.*

The limit of our model as $q \to \infty$ corresponds to the permutation model, and our result is consistent since the shortest cycle in a random permutation of n elements is asymptotically $e^{-\gamma}\log n$ (see [15]).

5. Distinct degree factorization

We estimate the probability that the distinct degree factorization is the full factorization. This is of interest for factorization algorithms as direct methods are known to compute such a distinct degree factorization.

THEOREM 5. *The probability that the irreducible factors of a random polynomial in $\mathbb{F}_q[x]$ of degree n be all of distinct degrees is asymptotically, as $n \to \infty$,*

$$e^{-\gamma_q} = \prod_n \frac{1 + I_n q^{-n}}{(1 - q^{-n})^{-I_n}},$$

and $\gamma_q \to \gamma$ (Euler's constant) as q tends to infinity. We have the following numerical values $e^{-\gamma_2} \approx 0.3967$, $e^{-\gamma_3} \approx 0.4693$, $e^{-\gamma_4} \approx 0.4983$, $e^{-\gamma_5} \approx 0.5137$, and $e^{-\gamma} \approx 0.5614$.

6. Conclusions

A large class of parameters relative to the irreducible factor decomposition of polynomials can clearly be studied by these elementary techniques. Amongst the relevant literature, we cite the results of Mignotte and Nicolas who proved that the degree of the splitting field of a random polynomial of degree n is "almost surely" close to $e^{\log^2 n}$. See [12, 13].

Bibliography

[1] Berlekamp (Elwyn R.). – *Algebraic Coding Theory.* – Mc Graw-Hill, 1968, revised 1984 edition.

[2] Car (Mireille). – Factorisation dans $\mathbb{F}_q[x]$. *Comptes–Rendus de l'Académie des Sciences*, vol. 294 (Ser. I), 1982, pp. 147–150.

[3] Comtet (L.). – *Advanced Combinatorics.* – Reidel, Dordrecht, 1974.

[4] Flajolet (Philippe) and Odlyzko (Andrew M.). – Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, vol. 3, n° 2, 1990, pp. 216–240.

[5] Flajolet (Philippe) and Soria (Michèle). – Gaussian limiting distributions for the number of components in combinatorial structures. *Journal of Combinatorial Theory, Series A*, vol. 53, 1990, pp. 165–182.

[6] Flajolet (Philippe) and Soria (Michèle). – General combinatorial schemas: Gaussian limit distributions and exponential tails. *Discrete Mathematics*, vol. 114, 1993, pp. 159–180.

[7] Gao (Zhicheng) and Richmond (L. Bruce). – Central and local limit theorems applied to asymptotic enumerations IV: Multivariate generating functions. *Journal of Computational and Applied Mathematics*, vol. 41, 1992, pp. 177–186.

[8] Goulden (Ian P.) and Jackson (David M.). – *Combinatorial Enumeration.* – John Wiley, New York, 1983.

[9] Greene (D. H.) and Knuth (D. E.). – *Mathematics for the analysis of algorithms.* – Birkhauser, Boston, 1982, 2nd edition.

[10] Knopfmacher (John) and Knopfmacher (Arnold). – Counting irreducible factors of polynomials over a finite field. *Discrete Mathematics*, vol. 112, 1993, pp. 103–118.

[11] Knuth (Donald E.). – *The Art of Computer Programming.* – Addison-Wesley, 1981, 2nd edition, vol. 2: Seminumerical Algorithms.

[12] Mignotte (M.) and Nicolas (J.-L.). – Statistiques sur $\mathbb{F}_q[x]$. *Annales de l'Institut Henri Poincaré, Série B*, vol. XIX, n° 2, 1983, pp. 113–121.

[13] Nicolas (J.-L.). – A Gaussian law on $\mathbb{F}_q[x]$. In *Topics in Classical Number Theory, Colloquia Mathematica Societatis Janos Bolyai*, vol. 34, pp. 1127–1162. – 1981.

[14] Odlyzko (A. M.). – Asymptotic enumeration methods. – Preprint, March 1993. To appear as a chapter in the *Handbook of Combinatorics*, R. Graham, M. Grötschel and L. Lovász, ed.

[15] Shepp (L. A.) and Lloyd (S. P.). – Ordered cycle lengths in a random permutation. *Transactions of the American Mathematical Society*, vol. 121, 1966, pp. 340–357.

The Cost Structure of Quadtrees

Bruno Salvy
INRIA Rocquencourt

October 4, 1993

[summary by Philippe Dumas and Michèle Soria]

Abstract

Many characteristics of quadtrees, like search costs and page occupancy, are precisely analysed. The mean value of such parameters are shown to have generating functions of hypergeometric type. Integral representations and singularity analysis give explicit forms for many structural constants of those trees.

The quadtree structure is a natural generalization of binary search trees to d-dimensional data. It constitutes a fundamental hierarchical representation of point data in higher dimensional spaces, which is used in many different fields like data bases or image processing (see e.g. Samet's book [7]).

Interesting parameters in the study of quadtrees are path length (related to the cost of searching or inserting data) and the page occupancy in the case of quadtrees depending on an integer parameter b representing a page capacity. These additive parameters have been largely investigated: for example, the average path length for d-dimensional quadtrees of size n is equivalent to $\frac{2}{d}\, n \log n$ [3]; the page occupancy, in the case $d = 2$, approaches 33% [5].

The method used by Flajolet *et alii* in [3, 4, 5] relies on studying a linear differential equation of order d and the local behaviour of its solutions. The results presented here rely on a new method of attack, in which the dimension d is only a parameter in some linear differential equation of order 1. With this method, it is possible to get more precise asymptotic expansions for the average value of parameters, and to obtain formal expressions for the coefficients of these expansions for all d's (but goldies). The previous method, in the case $d > 2$, only gave an equivalent and the involved constant was not always reachable.

1. Classical method

All cost functions which are studied here are additive and the mean value f_n over all d-quadtrees of size n satisfies a recurrence

$$f_n = t_n + 2^d \sum_{k=0}^{n} \pi_{n,k}\, f_k, \tag{1}$$

where t_n is a toll function, that is to say the cost of dividing into sub-problems and reconstructing from sub-problems. For example, if $t_n = 1$ then f_n is the number of nodes, and if $t_n = n$ then f_n is the average path length. Besides, $\pi_{n,k}$ is the probability that the first subtree has size k, which has value

$$\pi_{n,k} = \frac{1}{n} \sum_{k \le N_1 \le \cdots \le N_{d-1} \le n-1} \frac{1}{(N_1+1)\cdots(N_{d-1}+1)}.$$

Translating equation (1) into generating functions gives for $f(z) = \sum_{n=0}^{\infty} f_n\, z^n$ the integral equation

$$f(z) = t(z) + 2^d\, J^{d-1} I f(z), \tag{2}$$

where I and J are two linear operators defined by

$$If(z) = \int_0^z \frac{f(t)}{1-t}\, dt, \quad Jf(z) = \int_0^z \frac{f(t)}{t(1-t)}\, dt.$$

The functional equation (2) may be expressed as a linear differential equation of order d, namely

$$\left[z(1-z)\frac{d}{dz}\right]^d \{f(z) - t(z)\} = 2^d z f(z). \tag{3}$$

The asymptotic behaviour of the sequence (f_n) depends on the dominant singularities of $f(z)$, which is a solution of (3).

Let us remind that for a linear differential equation

$$a_k(z)\, y^{(k)}(z) + \cdots + a_0(z)\, y(z) = 0,$$

where the a_i's are polynomials, the singular points are at the roots α of $a_k(z)$. Moreover the local behaviour of the solution may have two forms

$$lbreg(z) = \left(1 - \frac{z}{\alpha}\right)^s \sum_{n \geq 0} P_{k,n}\left(\log(\alpha - z)\right) \left(1 - \frac{z}{\alpha}\right)^{n/q}$$

or

$$lbirreg(z) = lbreg(z) \times \exp\left[P\left(\frac{1}{(1 - z/\alpha)^{1/q}}\right)\right],$$

according to the regular or irregular type of the singular point α [9]. It is possible to compute all these quantities by a method of indeterminate coefficients [8]. In this way, one obtains a basis of singular solutions (the series may be convergent or divergent). But a basis is not sufficient and one must also find the coordinates of the studied solution with respect to the basis.

EXAMPLE. The generating function of the average path length in dimension $d = 2$ satisfies the linear differential equation [3]

$$z(1-z)^2 P''(z) + (1-2\,z)(1-z)P'(z) - 4\,P(z) = \frac{1+3\,z}{(1-z)^2}.$$

In this case, it is possible to give an explicit solution in the form of a hypergeometric function, but we neglect this point of view to illustrate the general method.

First we deal with the homogeneous equation. It has two singularities $\alpha = 0,\, 1$, and we find two solutions

$$f_1(z) = \frac{1}{(1-z)^2}\left[1 - \frac{2}{3}(1-z) + \cdots\right], \quad f_2(z) = (1-z)^2\left[1 + \frac{6}{5}(1-z)\cdots\right].$$

A particular local solution is

$$f_0(z) = \frac{1}{(1-z)^2}\left[\log\frac{1}{1-z} - \frac{2}{3} + \cdots\right].$$

So the general solution

$$f(z) = f_0(z) + c_1\, f_1(z) + c_2\, f_2(z)$$

has the singular behaviour

$$\frac{1}{(1-z)^2}\log\frac{1}{1-z}+\frac{c_1-2/3}{(1-z)^2}+\cdots,$$

hence for the coefficients

$$f_n = n\log n + (c_1+\gamma-5/3)\,n+\cdots. \tag{4}$$

Thus the recurrence gives first a linear differential equation, next the dominant singularity and local behaviour, and eventually the asymptotic behaviour of f_n. Alas we cannot compute the coefficient c_1, except numerically.

2. New method

The new method relies first on Euler transform, which yields a first order recurrence instead of a full history recurrence, and second on analytically continuation of Taylor series.

2.1. Euler transform. The Euler transform

$$f^*(z)=\frac{1}{1-z}f\left(\frac{-z}{1-z}\right)$$

is an involution and the associated relation on coefficients is

$$f_n=\sum_{k=0}^{n}(-1)^k\binom{n}{k}f_k^*.$$

To abbreviate, we note $Z=-z/(1-z)$. According to (2), the function $f^*(z)$ satisfies a new functional equation

$$(1-Z)f^*(Z)=(1-Z)t^*(Z)+2^dJ^{d-1}I(1-Z)f^*(Z),$$

which involves two operators, which are much simpler than the preceding ones,

$$I(1-Z)f^*(Z)=-\int_0^Z f^*(u)\,du,\quad\text{and}\quad Jg(Z)=\int_0^Z\frac{g(u)}{u}\,du.$$

Moreover the underlying recurrence is merely of order 1,

$$f_n^*=u_n+\left[1-\left(\frac{2}{n}\right)^d\right]f_{n-1}^* \tag{5}$$

($u_n=t_n^*-t_{n-1}^*$) and it is to be compared with the original recurrence

$$f_n=t_n+2^d\sum_{k=0}^{n}\pi_{n,k}f_k.$$

Example. For the average path length recurrence, (5) is easy to solve because $u_n=\delta_{n,2}-\delta_{n,1}$ is zero for $n\ge 3$, hence

$$f_n^*=\prod_{k=3}^{n}\left(1-\left(\frac{2}{k}\right)^d\right)\text{ for }n\ge 3.$$

As a result $f^*(z)$ is hypergeometric, hence $f(z)$ is hypergeometric too. More precisely we have

$$f(z)=\frac{z}{(1-z)^2}+\frac{z^2}{(1-z)^3}\,{}_{d+1}F_d\left(\begin{matrix}3-\omega_1,\ldots,3-\omega_d,1\\3,\ldots,3\end{matrix}\,\middle|\,z\right),$$

with the classical notation of generalized hypergeometric functions

$$ {}_pF_q\left(\begin{array}{c} a_1,\ldots,a_p \\ b_1,\ldots,b_q \end{array}\,\middle|\, z\right) = \sum_{n=0}^{\infty} \frac{(a_1)_n\cdots(a_p)_n}{(b_1)_n\ldots(b_q)_n}\frac{z^n}{n!}, $$

and $(a)_n = a(a+1)\cdots(a+n-1)$. The link between f_n and the hypergeometric series comes from the equality

$$ \prod_{k=3}^{n}\frac{k^d-2^d}{k^d} = \prod_{k}\prod_{\omega^d=2^d}\frac{k-\omega}{k} = \prod_{\omega}\frac{(3-\omega)(4-\omega)\cdots(n-\omega)}{3\cdot 4\cdots n}. $$

2.2. General theorem. In the preceding example, the expression of f_n is explicit. This is a general result because f_n^* satisfies

$$ f_n^* = A(n)\sum_{k=2}^{n}\frac{u_k}{A(k)}. $$

For paged trees, we have $u_k = (-1)^{k+1}\binom{n-2}{k-1}$ and for the number of leaves u_k is simply 1. We obtain in this way the following theorem [2].

THEOREM 1. *The expectation f_n of an additive cost function with toll t_n is*

$$ f_n = t_0 + n\left((2^d-1)t_0+t_1\right) + \sum_{k=2}^{n}(-1)^k\binom{n}{k}A_k\sum_{j=2}^{k}\frac{t_j^*-t_{j-1}^*}{A_j} $$

with

$$ A_n = \prod_{j=3}^{n}\left(1-\frac{2^j}{j^d}\right), \quad t_j^* = \sum_{k=0}^{j}(-1)^k\binom{j}{k}t_k. $$

But the asymptotic behaviour is not yet known. To get that damned behaviour we search for the behaviour of $f^*(z)$ at $-\infty$, since point $-\infty$ corresponds to point 1 (the dominant singularity of $f(z)$) through Euler transformation. Function $f^*(Z)$ is defined as a power series and we search for an integral representation, which permits an extension to the whole plane. The formula for analytic continuation of Taylor series [6] is

$$ g(-t) = \frac{1}{2i\pi}\int_{c-i\infty}^{c+i\infty}\varphi(-s)t^{-s}\frac{\pi}{\sin\pi s}\,ds $$

if we assume that

$$ g(t) = \sum_{n\geq 0}\varphi(n)\,(-t)^n $$

and $\varphi(s)$ is an analytical function, satisfying some growth conditions.

EXAMPLE. For the path length

$$ f_n^* = \prod_{k=3}^{n}\left(1-\left(\frac{2}{k}\right)^d\right) = \frac{A(n)}{A(2)} $$

with

$$ A(n) = \frac{1}{n(n-1)}\prod_{\substack{\omega^d=2^d\\ \omega\neq 2}}\frac{\Gamma(n+1-\omega)}{\Gamma(n+1)} $$

and we use

$$ f^*(-t) = \frac{1}{A(2)}\frac{1}{2i\pi}\int_{-3/2-i\infty}^{-3/2+i\infty}A(-s)t^{-s}\frac{\pi}{\sin\pi s}\,ds, $$

as an extension of our power series $f^*(-t)$. Next we shift the line of integration to the right and collect the residues. There is a pole of order 2 at $s=-1$ and the computation of the residue gives

$$f^*(-t) \underset{t\to\infty}{=} t + \frac{2}{d}\,t\left[\log t - 1 + \sum_{\omega} \psi(2-\omega) - \psi(2)\right] + O\left(\log t + t^{1-2\cos 2\pi/d}\right),$$

where $\psi = \Gamma'/\Gamma$. Using Euler transform, the preceding expansion, valid in a neighbourhood of $-\infty$, becomes an expansion in the neighbourhood of 1, which yields a more precise expression than the one obtained by the classical method –cf. (4)–

$$f_n = \frac{2}{d} n \log n + n\left(1 - \frac{1}{d} + \frac{2\,\gamma}{d} + \frac{2}{d} \sum_{\substack{\omega^d=2^d \\ \omega\neq 2}} [\psi(2-\omega) - \psi(2)]\right) + O\left(\log n + n^{1-2\,\cos 2\,\pi/d}\right).$$

EXAMPLE. For the number of leaves the formula is

$$f_n^* = A(n) \sum_{k=2}^{n} \frac{1}{A(k)}.$$

The problem is to make this expression an analytic function of n. A first way to do this is to use the series of differences (the term $n-1$ is a correction due to $\lim_k A(k) = 1$),

$$f_n^* = A(n)\left[\sum_{k\geq 2}\left(\frac{1}{A(k)} - \frac{1}{A(k+n-1)}\right) + n - 1\right].$$

A second way is to write

$$f_n^* = \lim_{u\to 1} A(n) \sum_{k\geq 2}\left(\frac{u^k}{A(k)} - \frac{u^{k+n-1}}{A(k+n-1)}\right).$$

In both cases we obtain

$$f_n = n\left[1 - \prod_{\substack{\omega^d=2^d \\ \omega\neq 2}} \Gamma(2-\omega)\left(1 + \sum_k \frac{A'(k)}{A(k)}\right)\right] + \ldots$$

For $d=2$, the factor of n has value $4\,\pi^2 - 39 \simeq 0.47841762$.

EXAMPLE. Eventually the page occupancy for quadtrees with page capacity b gives rise to the sequence

$$f_n^* = A(n) \sum_{k=b+1}^{n} \frac{\binom{k-2}{b-1}}{A(k)}.$$

The continuation of f_n^* as an analytical function of n is more subtle. The first way needs the b first terms of the asymptotic expansion of f_n^* to be precomputed, and this is not satisfactory. The second way uses

$$f_n^* = \lim_{v\to 1}\left\{A(n) \sum_{j\geq b+1}\left[\binom{j-2}{b-1}\frac{v^j}{A(j)} - \frac{\Gamma(j+n-2)}{(b-1)!\Gamma(j+n-b-1)}\frac{v^{j+n-1}}{A(j+n-1)}\right]\right.$$
$$\left. - A(n) \sum_{j=2}^{b} \frac{\Gamma(j+n-2)}{(b-1)!\Gamma(j+n-b-1)}\frac{v^{j+n-1}}{A(j+n-1)}\right\}.$$

It is noteworthy that the bounds of the sums are independent of n. So that it is possible to formally compute the constants in the asymptotic expansion given by the theorem. By the former method, these constants were attainable only by numerical computation.

3. Conclusion

This new method, which treats the dimension d as a parameter, permits to study precisely the additive characteristics of quadtrees: full asymptotic expansions are available, coefficients of these expansions are formally computable. But the computation is rather difficult and may involve summation of multiple series. Moreover, only additive parameters can be dealt with. (Such an important parameter as the height must be tackled with other methods [1].) Still this method has the advantage of generality and may be applied to a wide class of problems.

Bibliography

[1] Devroye (Luc) and Laforest (Louise). – An analysis of random d–dimensional quad trees. *SIAM Journal on Computing*, vol. 19, 1990, pp. 821–832.

[2] Flajolet (Ph.), Labelle (G.), Laforest (L.), and Salvy (B.). – *The Cost Structure of Quadtrees.* – Technical Report n° 2249, Institut National de Recherche en Informatique et en Automatique, April 1994.

[3] Flajolet (Philippe), Gonnet (Gaston), Puech (Claude), and Robson (J. M.). – The analysis of multi-dimensional searching in quad–trees. In *Proceedings of the Second Annual ACM–SIAM Symposium on Discrete Algorithms.* pp. 100–109. – Philadelphia, 1991.

[4] Flajolet (Philippe) and Lafforgue (Thomas). – Search costs in quadtrees and singularity perturbation asymptotics. *Discrete and Computational Geometry*, vol. 12, n° 4, 1994.

[5] Hoshi (Mamoru) and Flajolet (Philippe). – Page usage in a quadtree index. *BIT*, vol. 32, 1992, pp. 384–402.

[6] Lindelöf (Ernst). – Le calcul des résidus et ses applications à la théorie des fonctions. In *Collection de monographies sur la théorie des fonctions, publiée sous la direction de M. Émile Borel.* – Gauthier-Villars, Paris, 1905. Reprinted by J. Gabay, 1989.

[7] Samet (Hanan). – *The Design and Analysis of Spatial Data Structures.* – Addison–Wesley, 1990.

[8] Tournier (Évelyne). – *Solutions formelles d'équations différentielles.* – Doctorat d'État, Université scientifique, technologique et médicale de Grenoble, 1987.

[9] Wasow (W.). – *Asymptotic Expansions for Ordinary Differential Equations.* – Dover, 1987. A reprint of the John Wiley edition, 1965.

Ramanujan's Q-function and Computer Science Applications

Helmut Prodinger
Technical University of Vienna

October 25, 1993

[summary by Xavier Gourdon]

Abstract

The function $Q(n) = 1 + \frac{(n-1)}{n} + \frac{(n-1)(n-2)}{n^2} + \cdots$ plays a vital role in the analysis of several algorithms and some Discrete Mathematics problems, like the classical birthday problem, hashing with linear probing, etc. We exhibit some of those connections. Also, we discuss how an original question of Ramanujan can be attacked by methods from complex analysis.

1. Ramanujan's Q-function

Let n be a positive integer. We defined the Ramanujan's Q-function of n by

$$Q(n) = 1 + \left(\frac{n-1}{n}\right) + \left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right) + \cdots = \sum_{k=1}^{n} \frac{(n)_k}{n^k}.$$

2. A first problem: Hashing with linear probing

This was Knuth's first analysis. We consider the following game: n players arrive sequentially at a random chair in $\{1, \ldots, m\}$ with $m \geq n$. If the chair is already occupied, then the corresponding player try to sit on the next chair, and continue until the chair is free. We denote by $d(m, n)$ the average distance that the n-th player has to travel.

For example, consider 6 players A, B, C, D, E, F with 6 chairs in $\{1, 2, 3, 4, 5, 6\}$, and suppose

$$A \to 3, \quad B \to 1, \quad C \to 4, \quad D \to 3-4-5, \quad E \to 1-2, \quad F \to 3-4-5-6.$$

Players A, B, C travelled a distance 0, player D a distance 2, player E a distance 1 and player F a distance 3.

Our parameter of interest is

$$\delta(m, n) = \frac{d(m, 1) + \cdots + d(m, n)}{n}.$$

All m^n "hash sequences" are assumed to be equally likely. An analysis leads to

$$d(m, n) = \frac{(m-n)m^{1-n}}{2} \cdot \sum_{r \geq 0} r \binom{n-1}{r} (r+1)^r (m-r-1)^{n-r-2}.$$

Thanks to Abel's binomial theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x(x-kz)^{k-1}(y+kz)^{n-k},$$

we find

$$d(m,n) = \frac{1}{2}\left(2\frac{n-1}{m} + 3\frac{(n-1)(n-2)}{m^2} + \cdots\right)$$

so that

$$\delta(m,n) = \frac{1}{2}\left(\frac{n-1}{m} + \frac{(n-1)(n-2)}{m^2} + \cdots\right).$$

Asymptotic development of $\delta(m,m)$. We consider the case $m = n$, for which $\delta(m,n) = \frac{1}{2}(Q(n)-1)$. We now have to find the asymptotic expansion of $Q(n)$. We introduce

$$R(n) = 1 + \frac{n}{n+1} + \frac{n^2}{(n+1)(n+2)} + \cdots$$

We have the relation

$$Q(n) + R(n) = \frac{n!e^n}{n^n} \sim \sqrt{2\pi n} + \cdots$$

(the last term follows from Stirling's formula). The expression of $R(n)$ in terms of the incomplete gamma function $\gamma(a,x) = \int_0^x e^{-t}t^{a-1}\,dt$

$$R(n) = \frac{n!e^n}{n^n} \cdot \frac{\gamma(n,n)}{(n-1)!},$$

together with the asymptotic development [6, §1.2.11.3]

$$\frac{\gamma(x+1,x+y)}{\Gamma(x+1)} = \frac{1}{2} + \frac{y-2/3}{\sqrt{2\pi}}x^{-1/2} + \cdots \qquad (x \to \infty,\ y \text{ fixed})$$

leads to

$$R(n) = \sqrt{\frac{\pi n}{2}} + \frac{1}{3} + \frac{1}{12}\sqrt{\frac{\pi}{2n}} + \cdots \quad \text{and} \quad Q(n) = \sqrt{\frac{\pi n}{2}} - \frac{1}{3} + \frac{1}{12}\sqrt{\frac{\pi}{2n}} + \cdots$$

Another interesting proof can be found in [4, p. 5].

3. The Flajolet-Gardy-Thimonier approach

We consider the general birthday paradox. The year has m days, with probability of appearance $p_1, \ldots, p_m$. We denote by $E(B_j)$ the expected number of drawing until you first encounter j days with multiplicity at least k (the classical case is $k = 2$, $j = 1$, with $m = 365$). The problem can be coded by an appropriate language, which translates easily in terms of exponential generating functions, and using the Laplace-Borel transform to get back to ordinary generating functions (see [3]). We have

$$E(B_j) = \sum_{q=0}^{j-1} \int_0^\infty [u^q] \prod_{i=1}^{m} \left(e_{k-1}(p_i z) + u(e^{p_i z} - e_{k-1}(p_i z))\right) e^{-z}\,dz,$$

where $e_{k-1}(x) = \sum_{j=0}^{k-1} x^j/j!$. Let's see special instances. When $j = 1$ (first k-hit), we have

$$E(B_1) = \int_0^\infty \left(\prod_{i=1}^{\infty} e_{k-1}(p_i z)\right) e^{-z}\,dz,$$

and for the equiprobable case $p_i = 1/m$

$$E(B_1) = \int_0^\infty \left(e_{k-1}(\frac{z}{m})\right)^m e^{-z}\,dz,$$

so that for $k = 2$

$$E(B_1) = \int_0^\infty \left(1 + \frac{z}{m}\right)^m e^{-z}\, dz = 1 + Q(m). \tag{1}$$

For $k = 1$ and $j = m$ (coupon collector's problem)

$$E(B_m) = \int_0^\infty \left(1 - \prod_{i=1}^{m}(1 - e^{-p_i z})\right) dz$$

so that in the equiprobable case

$$E(B_m) = m \left(\sum_{q=1}^{m}(-1)^{q-1}\binom{m}{q}\frac{1}{q}\right) = mH_m.$$

4. A conjecture of Ramanujan about the Q-function

Define θ by

$$\sum_{k=0}^{n-1}\frac{n^k}{k!} + \frac{n^n}{n!}\theta = \frac{e^n}{2}.$$

It follows from [1, p. 181, entry 48] that $\theta \approx \dfrac{4+15n}{8+45n}$, and Ramanujan conjectured that always $\frac{1}{3} \le \theta \le \frac{1}{2}$. This was proved independently by Watson and Szegö. A more refined conjecture (also posed by Ramanujan) was proved only recently by Flajolet, Grabner, Kirschenhofer and Prodinger [4]. We have

$$\theta = \frac{1}{3} + \frac{4}{135(n+k)} \quad \text{where always} \quad \frac{2}{21} \le k \le \frac{8}{45}.$$

All this is related to the Q-function because

$$\frac{n^n}{n!}Q(n) = \frac{n^{n-1}}{(n-1)!} + \frac{n^{n-2}}{(n-2)!} + \cdots + 1 = \frac{1}{2}e^n - \theta\frac{n^n}{n!}.$$

Since

$$Q(n) + R(n) = \frac{n!e^n}{n^n}, \quad \text{we have} \quad \theta(n) = \frac{1}{2}(R(n) - Q(n)),$$

and the Ramanujan's problem can be rephrased as

$$D(n) = R(n) - Q(n) = \frac{2}{3} + \frac{8}{135(n+k)} \quad \text{where always} \quad \frac{2}{21} \le k \le \frac{8}{45}.$$

The approaches followed by Ramanujan himself and later authors all make use of the integral (1) and proceed using the Laplace method for the asymptotic evaluation of integrals. The idea used in [4] is rather different and uses complex analysis applied to famous Knuth's "tree function" $y(z)$, defined implicitly by the equation

$$y(z) = ze^{y(z)}.$$

Thanks to Lagrange's inversion formula, we get

$$[z^n]y(z) = \frac{n^{n-1}}{n!}, \qquad Q(n)\,\frac{n^{n-1}}{n!} = [z^n]\log\frac{1}{1-y(z)}$$

and

$$D(n)\,\frac{n^{n-1}}{n!} = [z^n]\log\frac{(1-y(z))^2}{2(1-y(z)e^{1-y(z)})}.$$

By singularity analysis [5] this equality can be used to get easily the first 10-terms of the asymptotic expansion of $D(n)$:

$$D(n) = \frac{2}{3} + \frac{8}{135n} + \cdots - \frac{479}{561330n^9} + O\left(\frac{1}{n^{10}}\right).$$

An estimation of the remainder was done by saddle point heuristic in the y-plane, and it is proved in [4] that this remainder is less than $10^{-7}/n^3$ for $n \geq 116$. Then one can show the desired bounds for $n \geq 116$ and check the remaining 115 cases "by hand".

5. More of Knuth's wisdom

The function $Q(n)$ shows up in unexpected places. For example, Cauchy [2, pp. 62-73] proved that

$$\frac{1}{n^n} \sum_k \binom{n}{k} k^k (n-k)^{n-k} = 1 + Q(n).$$

A Q-algebra. It is possible to develop a Q-algebra theory [7, p. 190]. For every sequence $(a_1, a_2, \dots)$ and every positive integer n, we consider

$$Q(a_1, a_2, \dots ; n) = \sum_{k \geq 1} a_k \frac{(n)_k}{n^k}.$$

For example, $Q(1, 1, \dots ; n) = Q(n)$. We have the two identities

$$rQ(a_1, a_2, \dots) + sQ(b_1, b_2, \dots) = Q(ra_1 + sb_1, ra_2 + sb_2, \dots)$$

and

$$Q(a_1, 2a_2, 3a_3, \dots ; n) = nQ(a_1, a_2 - a_1, a_3 - a_2, \dots ; n)$$

(for the latest, write $k = n - (n-k)$ — like Abel's partial summation). Therefore

- $Q(1, 2, 3, \dots ; n) = nQ(1, 0, 0, \dots ; n) = n$,
- $Q(1^2, 2^2, 3^2, \dots ; n) = nQ(1, 1, 1, \dots ; n) = nQ(n)$,
- $Q(1^3, 2^3, 3^3, \dots ; n) = 2n^2 - nQ(n)$,
- $Q(1^4, 2^4, 3^4, \dots ; n) = 3n^2 Q(n) - 3n^2 + nQ(n)$.

We can always write such a formula for $Q(1^k, 2^k, 3^k, \dots)$, with certain coefficients. They can be arranged in a triangle. The inverse triangle is also of interest, and the coefficients have combinatorial interpretations in terms of permutations of certain multisets.

Bibliography

[1] Berndt (Bruce C.). – *Ramanujan's Notebooks, Part II.* – Springer Verlag, 1989.

[2] Cauchy (A.). – *Exercices de Mathématiques.* – Paris, 1826.

[3] Flajolet (Philippe), Gardy (Danièle), and Thimonier (Loÿs). – Birthday paradox, coupon collectors, caching algorithms, and self–organizing search. *Discrete Applied Mathematics*, vol. 42, 1992.

[4] Flajolet (Philippe), Grabner (Peter), Kirschenhofer (Peter), and Prodinger (Helmut). – *On Ramanujan's Q–function.* – Research Report n° 1760, Institut National de Recherche en Informatique et en Automatique, October 1992. 13 pages. To appear in the *Journal of Computational and Applied Mathematics.*

[5] Flajolet (Philippe) and Odlyzko (Andrew M.). – Singularity analysis of generating functions. *SIAM Journal on Discrete Mathematics*, vol. 3, n° 2, 1990, pp. 216–240.

[6] Knuth (Donald E.). – *The Art of Computer Programming.* – Addison-Wesley, 1968, vol. 1: Fundamental Algorithms. Second edition, 1973.

[7] Knuth (Donald E.). – An analysis of optimal caching. *Journal of Algorithms*, vol. 6, 1985, pp. 181–199.

Sizes of Relations: a Dynamic Analysis

Danièle Gardy
Université de Versailles Saint-Quentin

February 7, 1994

[summary by Dominique Gouyou-Beauchamps]

1. Introduction

Among the parameters that can be defined on relational databases, the sizes of the relations, either present in the database or computed by application of a relational operator, have long been recognized as important parameters in query optimization.

The basic objects we[1] consider are *relations*, which are sets of (distinct) tuples. They can be seen as tables: a row represents a tuple, and the number of lines is the number of elements of the relation (its *size*); the columns are called the *attributes*. The *projection* of a relation on a subset of the set of attributes is a new relation, obtained by suppressing the corresponding columns, then all the duplicated rows in the resulting table: We keep only one instance of each tuple. We give in Figure 1 an instance of a relation $R[X, Y]$ and its projection denoted $\pi_X(R)$ on the attribute X.

Let us now consider two relations $R[X, Y]$ and $S[X, Z]$ with a common attribute X. The *equijoin* of R and S on their common attribute X, denoted $R \bowtie S$, has three attributes X, Y and Z; it is composed of all triples (x, y, z) such that (x, y) belongs to R and (x, z) belongs to S (see again Figure 1).

The *semijoin* is another operator on relations; although it can be defined using the projection and equijoin: $R[X, Y] \rhd S[X, Z] = \pi_{XY}(R \bowtie S) = R \bowtie \pi_X(S)$ (Figure 1 presents instances of $R \rhd S$ and $S \rhd R$).

We use the terms *initial relation* for the relations to which we apply a relational algebra operator (projection, equijoin or semijoin), and *derived relation* for the relation resulting from the operation.

We gave in former papers [4, 5] conditions which ensure that, in the static cases (i.e., at a given time), the size of a derived relation, obtained by a projection, an equijoin or a semijoin, follows a normal limiting distribution. Our goal here is to extend these results when the database is submitted to a sequence of queries and updates.

In the static case, we study the conditional distribution of the sizes of the derived relations obtained by a projection or by a join, assuming that the sizes of the initial relations are known. In the dynamic case, we want to study the influence of updates and queries on the sizes of initial relations and derived relations. To this effect, we shall use a modelization in terms of urn models.

2. Urn models and databases

We consider a sequence of d urns, each urn being labelled with a distinct value of the attribute X. To each tuple of the relation R, we associate a ball labelled by the value of the tuple on the column X; this ball falls into the corresponding urn. An equivalent way of seeing this phenomenon

[1] The original articles by D. Gardy and G. Louchard can be found in [6, 7].

R

X	Y
x_0	y_0
x_0	y_1
x_1	y_2
x_2	y_3

S

X	Z
x_0	z_0
x_0	z_1
x_1	z_1
x_3	z_2

$\pi_X(R)$

X
x_0
x_1
x_2

$R \bowtie S$

X	Y	Z
x_0	y_0	z_0
x_0	y_0	z_1
x_0	y_1	z_0
x_0	y_1	z_1
x_1	y_2	z_1

$R \rhd S$

X	Y
x_0	y_0
x_0	y_1
x_1	y_2

$S \rhd R$

X	Z
x_0	z_0
x_0	z_1
x_1	z_1

FIGURE 1. Examples of relations $R[X, Y]$ and $S[X,Z]$ with the projection $\pi_X(R)$ of R on X, the equijoin of R and S on X and the semijoins of R and S, and S and R.

is to consider instead that we have a finite supply of balls, and that we allocate them at random among the d urns, each trial being independent of the others. Each ball then receives the label of the urn it falls into. After coupling all the tuples of the initial relation R with urns, some urns are empty and some contain at least one ball. The number of urns with at least one ball is exactly the number of tuples in the projection of the relation R.

In the rest of the paper, we shall use indifferently the terms *relation size* and *number of balls* or *number of tuples*, and the terms *projection size* and *number of non-empty urns*.

3. Static models

Here, we consider the case of a *free* relation, i.e., there is total independence between the values taken by different tuples. In this framework, we obtain the following theorems [4, 8, 9].

THEOREM 1 ([4]). *Let $R[X,Y]$ be a free relation with a uniform probability distribution on the domain of attribute X. Then the probability distribution of the size of the projection of R on attribute X, conditioned by the size $r = Ad_X + o(d_X)$ of relation R with A a positive constant is asymptotically normal when $d_X \to \infty$. The asymptotic mean and variance are given by $\mu = \mu_0 d_X$ and $\sigma^2 = \sigma_0^2 d_X$ where μ_0 and σ_0^2 are constants that depend on the probability distribution on attribute Y.*

We consider relations where attribute X is a *key*, i.e., in a given instance of the relation the x-value of a tuple uniquely determines its y-value.

THEOREM 2 ([4]). *Let $R[X,Y]$ be a relation with a key X, and $S[X,Z]$ a relation with a key U. We assume that the probability distribution on D_X (the finite domain on which a probability distribution is defined for X) is uniform. The probability distributions on D_Y and D_Z are arbitrary. The sizes r and s of the relations R and S are assumed to satisfy $r < d_X$, $d_X = o(r^{3/2})$, and $s = Bd_X(1 + o(1))$. Then the probability distribution of the size of the semijoin of R and S on attribute X, conditioned by the sizes of R and S, is asymptotically normal. The mean and variance have for asymptotic values $\mu = (1 - e^{-B})r$ and $\sigma^2 = r((e^B - 1)/e^{2B} - rB/d_X e^{2B})$.*

We denote $p_{i,d}$ the probability that the ith element of the domain is selected when choosing at random an element of a finite domain D of size d. Let $\lambda_R(t) = \prod_{1 \le i \le d}((1 + p_{i,d}t)/(1 + p_{i,d}))$ be the

generating function associated with the probabilities of the set of distinct items for relation R. We assume that $\lambda_R(t) \neq (1+t)/2$.

THEOREM 3 ([9, 8]). *Let $R[X,Y]$ (resp. $S[X,Z]$) be a relation with a key Y (resp. Z). The sizes r and s of the relations R and S are assumed to satisfy $r = Ad_X + o(d_X)$ and $s = Bd_X + o(d_X)$. Let $g_R(y)$ (resp. $g_S(z)$) be the function $g_R(y) = y\frac{\lambda_R'}{\lambda_R}(y)$ (resp. $g_S(z) = z\frac{\lambda_S'}{\lambda_S}(z)$). Constants A and B are such that: $lim_{y\to+\infty} g_R(y) > A$ and $lim_{z\to+\infty} g_S(z) > B$. We assume that the probability distribution on D_X is uniform. Then the probability distribution of the size of the equijoin of R and S on attribute X, is asymptotically normal when $d_X \to \infty$. The mean and variance have for asymptotic values $\mu = \sigma^2 = \frac{rs}{d_X} \approx ABd_X$.*

4. Dynamic models

We shall denote by $p_{\mathcal{I}}$, $p_{\mathcal{D}}$ and $p_{\mathcal{Q}}$ the probability of making an insertion, a deletion or a query. We can choose non-equal probabilities for insertion and deletion, as long as the probability of an insertion is at least equal to the probability of a deletion: $p_{\mathcal{I}} \geq p_{\mathcal{D}}$.

If we choose to perform a deletion, the conditional probability of deleting a given ball is $1/n$, n being the number of balls at this time.

Assuming that the urn size is infinite corresponds, in terms of relational database, to a relation with a key on the attribute suppressed in the projection. As we also want to study relations without keys, we need to extend the models to the case where *the urns have a finite capacity* (there are δ places for balls). If we choose to perform an insertion, we must give the conditional probability of inserting a ball into an urn, and this is the place where the infinite and finite models differ. In the infinite urn model, each urn has the same probability of getting the new ball: $1/d$, with d the number of urns. In the finite urn model, we can view each urn as a collection of δ distinguishable cells, and each empty cell, whatever the urn it belongs to, has the same conditional probability of receiving the ball, knowing that we have chosen to perform an insertion. Thus the probability that we put a ball in urn V_i is $v_i/(d\delta - n)$ where v_i is the number of empty cells in V_i and n is the number of balls at this time.

We denote by $\Rightarrow$ the weak convergence of random function in the space of all right-continuous functions having left limits and endowed with the Skorohod metric (see Billingsley [1]). All convergences will be defined for $n \to +\infty$.

We study two related stochastic processes, describing respectively the number of balls denoted by $\mathcal{P}$, and size of the projection (number of non-empty urns), denoted by $\mathcal{Q}$; we shall show that each of these processes has a deterministic component of order n, and a random component of order $\sqrt{n}$.

Let W be the number of balls at some time. We might choose the current number of steps (number of queries or updates) as a measure for the time, which would then belong to the interval $[0, 2n]$. However, we shall study the asymptotic behaviour of W when the time goes to infinity, and it is interesting to change the time scale by choosing a time nt for $t \in [0,2]$, and to normalize the random variable W. For all the models presented below, the number of tuples W has an expectation and a variance of order n, and we can show that, for a suitable function f_1 related to the type of process, and assuming that we start from an empty structure at time 0:

$$\frac{W([nt]) - nf_1(t)}{\sqrt{n}} \Rightarrow X(t), \qquad 0 \leq t \leq 2,$$

where the process X(t) is a Markovian Gaussian process whose covariance is denoted $f_2(s,t)$, $s \leq t$ [6].

THEOREM 4 ([6]). *The size $S([nt])$ of the projection at time nt is asymptotically a non-Markovian Gaussian process such that*

$$\begin{aligned} S([nt]) &\sim nG(t) + \sqrt{n}X_1(t), \\ E[S([nt])] &\sim nG(t), \\ \mathrm{Var}[S([nt])] &\sim n\Phi(t), \end{aligned}$$

where $X_1(t)$ is a non-Markovian Gaussian process whose covariance is denoted $\Psi_R(s,t)$ and where the functions G, Φ and Ψ_R can be given explicitly and depend on urn models (infinite or bounded) and on functions $f_1(t)$ and $f_2(s,t)$. The relative error in the density is $O(1/\sqrt{n})$.

5. Example

The processes can be divided in two families:

(1) the *weighted structure* in the sense of Flajolet *et al.* [3], Louchard [10], with a possibility function given by $pos(\mathcal{D}) = k$ for a k-size structure (there are k ways of deleting an element in a structure composed from k elements!);
(2) the classical *unweighted structure*.

We study *the unweighted structure family* and we consider updates $(\mathcal{I}+\mathcal{D})$ and queries $(\mathcal{Q})$ with arrival at a relation of size $2n\overline{x} + s\sqrt{n}$ at the time $2n$. We assume that we start from an empty structure at time 0. The mean and variance corresponding to one step are given by

$$\overline{x} = p_{\mathcal{I}} - p_{\mathcal{D}}, \qquad \sigma^2 = p_{\mathcal{I}} + p_{\mathcal{D}} - \overline{x}^2$$

and

$$\frac{W([nt]) - n\overline{x}t}{\sqrt{n}} \Rightarrow \sigma BB(t) + \frac{at}{2},$$

with BB a Brownian Bridge. The expectation and covariance are given by

$$f_1(t) = \overline{x}t + \frac{at}{2\sqrt{n}}, \qquad f_2(s,t) = \sigma^2\frac{s(2-t)}{2}, \qquad s \le t.$$

Now, if we assume that the urns have an infinite capacity, we get for the process *size of the projection*:

$$\begin{aligned} S([nt]) &\sim nG(t) + \sqrt{n}X(t), \\ G(t) &= \alpha(1 - e^{-\overline{x}t/\alpha}) + \frac{a}{\sqrt{n}}te^{-\overline{x}t/\alpha} \end{aligned}$$

where $X(t)$ is a non-Markovian Gaussian process whose covariance is

$$\Psi_R(t_1,t_2) = e^{-\overline{x}(t_1+t_2)/\alpha}\left[\alpha\left(e^{\frac{\overline{x}t_1}{\alpha}\left(\frac{t_1}{t_2}\right)^{p_{\mathcal{D}}/\overline{x}}} - 1\right) - \overline{x}t_1\left(\frac{t_1}{t_2}\right)^{p_{\mathcal{D}}/\overline{x}} + \sigma^2\frac{t_1(2-t_2)}{2}\right],$$

where $\alpha = \frac{d}{n}$.

6. Sketch of the proof

The first step is to study the process $\mathcal{P}$ describing the number of tuples in the initial relation. In the cases we are interested in, $\mathcal{P}$ happens to be a Gaussian process with a deterministic part $\mathcal{P}_0$, on which is superimposed a random part $\mathcal{P}_1$:

$$\mathcal{P} = \mathcal{P}_0 + \mathcal{P}_1.$$

The process $\mathcal{P}_0$ follows a deterministic curve $nf_1(t)$; the function f_1 is the expectation of the number of balls (or tuples in the initial relation), and the process $\mathcal{P}_1$ is a Markovian Gaussian process of order $\sqrt{n}$.

The process $\mathcal{P}$: *number of tuples* determines another process $\mathcal{Q}$: *size of the projection*. Before considering $\mathcal{Q}$, we shall study another process $\mathcal{Q}_0$, defined as the size of the projection of a relation R, when the size of R is given by the process $\mathcal{P}_0$ (which is a first-order approximation of $\mathcal{P}$). To this effect, we define two random variables, say Y_1 and Y_2, which are simply the size of the projection at different times t_1 and t_2. The covariance $\mathrm{Cov}(Y_1, Y_2)$ will allow us to characterize $\mathcal{Q}_0$ as a process composed of a deterministic part $G(t)$ and a random part $\sqrt{n}V(t)$.

We then consider the process $\mathcal{P}$ obtained by superimposing $\mathcal{P}_1$ on $\mathcal{P}_0$. We can again define two random variables *size of the projection* at the times t_1 and t_2; let us call them S_1 and S_2. It is possible to write their covariance as

$$\mathrm{Cov}(S_1, S_2) = \mathrm{Cov}(Y_1, Y_2) + \gamma(t_1)\gamma(t_2)f_2(t_1, t_2)$$

for a suitable function $\gamma(t)$, $f_2(t_1, t_2)$ being the covariance of the process $\mathcal{P}_1$ taken at different times t_1 and t_2. The covariance of Y_1 and Y_2 thus characterizes the "static" part, and the term added to it to get the covariance of S_1 and S_2 comes from the fact that the number of tuples $\mathcal{P}$ is itself a Gaussian process.

Once we have the covariance of the sizes of the derived relation at times t_1 and t_2, the next part is to show that the final process $\mathcal{Q}$ is still asymptotically a Gaussian process. More precisely, we show that $\mathcal{Q}$ has a part $\mathcal{Q}_0$ of order n coming from $\mathcal{P}_0$, on which is added a random part $\mathcal{Q}_1$ of order $\sqrt{n}$ coming from $\mathcal{P}_0$ and from $\mathcal{P}_1$:

$$\mathcal{Q} = \mathcal{Q}_0 + \mathcal{Q}_1.$$

6.1. $\mathrm{Cov}(Y_1, Y_2)$ **for a non-random static structure.** For each urn model (bounded or unbounded), there exist two functions $F(x)$ and $\Psi_{NR}(t_1, t_2)$ such that, if we consider the size of the projection of a relation, itself of size $nf_1(t)$, the asymptotic values of its expectation at time t_1, $E(Y_1)$, and of its covariance at distinct times t_1 and t_2, $\mathrm{Cov}(Y_1, Y_2)$, are:

$$E(Y_1) \sim nF(f_1(t)),$$
$$\mathrm{Cov}(Y_1, Y_2) \sim n\Psi_{NR}(t_1, t_2).$$

6.1.1. Unbounded urns

The one ball survival probability between t_1 and t_2 is denoted by $ps_{1,2}(t_1, t_2)$ ($= 1$ if $t_1 = t_2$).

Then, we obtain with $\mathcal{P}_0$ a deterministic process

$$E(\mathcal{P}_0) = nf_1(t),$$
$$F(X) = \alpha\left(1 - e^{-X/\alpha}\right),$$
$$\Psi_{NR}(t_1, t_2) = \alpha e^{-\frac{f_1(t_1)}{\alpha}}\left(e^{-\frac{f_3(t_1,t_2)}{\alpha}} - e^{-\frac{f_1(t_2)}{\alpha}}\right) - f_1(t_1)ps_{1,2}e^{-\frac{f_1(t_1)+f_1(t_2)}{\alpha}}.$$

6.1.2. Bounded urns

We can view the content of an urn with $\nu(t)$ balls as a population of $\nu(t)$ type 1 (balls) individuals and $\delta - \nu(t)$ type 2 (empty places) individuals. Let

$$p_{i,j}(t_1, t_2) = \Pr[\text{individual of type } i \text{ at time } t_1 \text{ is of type } j \text{ at time } t_2].$$

Then, we obtain:

$$F(X) = \alpha \left(1 - \left(1 - \frac{X}{\beta}\right)^{\sigma}\right),$$

$$\Psi_{NR}(t_1, t_2) = \left(1 - \frac{f_1(t_1)}{\beta}\right)^{\delta} \left[\alpha p_{2,2}^{\delta} - \left(1 - \frac{f_1(t_2)}{\beta}\right)^{\delta - 1} \left(\alpha p_{2,2} + (1 - \frac{1}{\delta}) f_1(t_1) f_7\right)\right],$$

where $\beta = \alpha\delta$.

6.2. $\mathrm{Cov}(S_1, S_2)$**.** Y_1 and Y_2 denote the size of the projection of a relation R at the times t_1 and t_2, when the number of tuples of R is given by the process $\mathcal{P}_0$. S_1 and S_2 denote the same quantities when the number of tuples of R is given by the process $\mathcal{P}$. We first compute the variation of $\mathrm{Cov}(Y_1, Y_2)$ introduced by assuming that the numbers of tuples are no longer fixed, but Gaussian random variables; this gives $n\Psi_C(t_1, t_2)$. Then we compute the actual covariance of S_1 and S_2 and we show that it is of the type $n\Psi_R(t_1, t_2)$; we also prove that the size of the projection is then a Gaussian process.

As the process $\mathcal{P}$ is obtained by adding a process $\mathcal{P}_1$ of order $\sqrt{n}$ to the process $\mathcal{P}_0$, itself of order n, the number n_1 of balls at time t_1 is given by:

$$n_1 = n \left(f_1(t_1) + \frac{\theta_1}{\sqrt{n}}\right) + O(1),$$

where θ_1 is a Gaussian random variable with mean 0 and covariance $f_2(s, t)$.

Setting $\gamma(t) = F'(f_1(t))$, we obtain:

$$E[Y_1] \sim n \left(F(f_1(t_1)) + \frac{\theta_1}{\sqrt{n}} \gamma(t_1)\right),$$

$$\Psi_C(t_1, t_2) = \Psi_{NR}(t_1, t_2) + \overline{\varphi}_1(t_1, t_2) \frac{\theta_1}{\sqrt{n}} + \overline{\varphi}_2(t_1, t_2) \frac{\theta_2}{\sqrt{n}} + O\left(\frac{1}{n}\right),$$

for some $\overline{\varphi}_1$ and $\overline{\varphi}_2$.

We know from previous work [4] that for a known size of the initial relation R at times t_1 and t_2 (static case), the projection size Y_1 and Y_2 are asymptotically Gaussian. Then for any ξ_1 and ξ_2,

$$E\left[e^{i(\xi_1 Y_1 + \xi_2 Y_2)}\right] \sim E\left(\exp[i(\xi_1 E[Y_1] + \xi_2 E[Y_2]) - \frac{1}{2}(\xi_1^2 \sigma^2(Y_1) + 2\xi_1 \xi_2 \,\mathrm{Cov}(Y_1, Y_2) + \xi_2^2 \sigma^2(Y_2))]\right).$$

Plugging the modified values for $E[Y_1]$ and $E[Y_2]$ into this equation, and substituting $\mathrm{Cov}(Y_1, Y_2)$ by $n\Psi_C(t_1, t_2)$ (and similarly for $\sigma^2(Y_1)$ and $\sigma^2(Y_2)$), we obtain:

$$E\left[e^{i(\xi_1 S_1 + \xi_2 S_2)}\right] \sim e^{A(t_1, t_2)} E\left[e^{B(t_1, t_2)}\right],$$

where the term $B(t_1, t_2)$ contains all the contribution from the Gaussian random variables θ_1 and θ_2 and is of the form $B(t_1, t_2) = i(\zeta_1 \theta_1 + \zeta_2 \theta_2)$.

This leads to:

$$E\left[e^{i(\xi_1 S_1+\xi_2 S_2)}\right] \sim \exp(i(\xi_1 nG(t_1)+\xi_2 nG(t_2)) - \frac{1}{2}(\xi_1^2 n\Psi_R(t_1,t_2)+2\xi_1\xi_2 n\Psi_R(t_1,t_2)+\xi_2^2 n\Psi_R(t_1,t_2)) + \textit{cubic terms in } \xi_1,\ \xi_2 + O\left(\frac{1}{\sqrt{n}}\right)).$$

Now remember that we are actually interested in the normalized process $S'([nt]) = (S([nt]) - nG(t))/\sqrt{n}$. Substituting ξ_1 by $\xi_1'/\sqrt{n}$ and ξ_2 by $\xi_2'/\sqrt{n}$, we get:

$$E\left[e^{i(\xi_1' S_1'+\xi_2' S_2')}\right] \sim \exp\left[-\frac{1}{2}(\xi_1'^2\Psi_R(t_1,t_2)+2\xi_1'\xi_2'\Psi_R(t_1,t_2)+\xi_2'^2\Psi_R(t_1,t_2)+O\left(\frac{1}{\sqrt{n}}\right)\right].$$

Then we state a new version, more precise, of theorem 4:

THEOREM 5 ([6]). *In the projection model, the size $S([nt])$ of the projection at time nt is asymptotically a non-Markovian Gaussian process with*

$$\begin{aligned}
S([nt]) &\sim nG(t)+\sqrt{n}X_1(t), \\
E[S([nt])] &\sim nG(t), && \textit{with} \quad G(t)=F(f_1(t)), \\
\mathrm{Cov}(S([nt_1]),S([nt_2])) &\sim n\Psi_R(t_1,t_2), && \textit{with} \quad \Psi_R(t_1,t_2)=\Psi_{NR}(t_1,t_2)+f_2(t_1,t_2)\gamma(t_1)\gamma(t_2), \\
\mathrm{Var}[S([nt])] &\sim n\Phi(t), && \textit{with} \quad \Phi(t)=\Psi_R(t,t)=\Psi_{NR}(t,t)+\gamma^2(t)f_2(t,t),
\end{aligned}$$

where $X_1(t)$ is a non-Markovian Gaussian process whose covariance is denoted $\Psi_R(s,t)$ and where the functions G, Φ and Ψ_R can be given explicitly and depend of urn models (infinite or bounded) and on functions $f_1(t)$ and $f_2(s,t)$. The relative error in the density is $O(1/\sqrt{n})$.

7. Projection maximum

We have shown that the process $\mathcal{Q}$ happens to be a Gaussian process $X(t)$ superimposed on a deterministic process $G(t)$:

$$S([nt]) = G(t)+X(t).$$

If we look for its maximum $m=\max\{G(t)+X(t)\}$, and for the time $t^\star$ at which this maximum occurs, it is equivalent to searching for the hitting time of $X(t)$ to the absorbing boundary $m-G(t)$. A theorem of Daniels [2] leads to the result.

Bibliography

[1] Billinsley (Patrick). – *Convergence of Probability Measures.* – John Wiley & Sons, 1968.

[2] Daniels (H. E.). – The maximum of a Gaussian process whose mean path has a maximum, with an application to the strength of bundles of fibres. *Advances in Applied Probability*, vol. 21, 1989, pp. 315–333.

[3] Flajolet (P.), Puech (C.), and Vuillemin (J.). – The analysis of simple list structures. *Information Sciences*, vol. 38, 1986, pp. 121–146.

[4] Gardy (D.). – Normal limiting distribution for projection and semijoin sizes. *SIAM Journal on Discrete Mathematics*, vol. 5, n° 2, 1992, pp. 219–248.

[5] Gardy (D.). – Join sizes, urn models and normal limiting distributions. *Theoretical Computer Science*, vol. 131, n° 2, August 1994.

[6] Gardy (D.) and Louchard (G.). – *Dynamic analysis of some relational data base parameters. I: Projections.* – Research Report n° 94-6, PRISM, University of Versailles, February 1994.

[7] Gardy (D.) and Louchard (G.). – *Dynamic analysis of some relational data base parameters. II: Equijoins and semijoins.* – Research Report n° 94-7, PRISM, University of Versailles, February 1994.

[8] Gardy (D.) and Puech (C.). – On the effect of joins operations on relation sizes. *ACM Transactions on Database Systems*, vol. 14, n° 4, 1989, pp. 574–603.

[9] Gardy (Danièle). – *Bases de données et allocations aléatoires: Quelques analyses de performance.* – Doctorate in Sciences, Université de Paris-Sud, 1989.

[10] Louchard (G.). – Random walks, Gaussian processes and list structures. *Theoretical Computer Science*, vol. 53, 1987, pp. 99–124.

Data Base Parameters: Equijoin and Semijoin

Guy Louchard
Département d'Informatique, Université Libre de Bruxelles

February 7, 1994

[summary by Danièle Gardy]

1. Introduction

This talk is a sequel to the talk *Sizes of relations: a dynamic analysis* by D. Gardy, and we shall make references to its summary [4]. A presentation of the database problem and of the model was given there, along with the dynamic study of the size of a relation obtained by *projection* of an initial relation. This second talk is centred around *joins*, i.e. operations building a "derived" relation from two initial relations. As was the case for projections, the joins can be described by urn models [1]. Although the computations become quite involved, due to the fact that we deal with bi-dimensional processes, the techniques are similar to those used for the projection.

This second talk begins by presenting in detail some points introduced in [4], such as the birth and death processes describing the number of balls in an urn, then turns to join models. The complete demonstrations are given in the full papers [2] (for the projections) and [3] (for the joins).

2. Urn models and processes

2.1. Urn models. We can describe a relation present in the database by an urn model, and its size by a random allocation model counting the number of balls allocated according to a certain scheme (see [4] in these proceedings). Now the equi- or semi-join of two relations R and S can be modelled in a similar way: Let X be the join attribute, and d the number of values X can take; we consider d distinguishable urns labelled by these values.

- We throw r red balls for the relation R, and s blue balls for the relation S, according to the rules reflecting the constraints on these two relations;
- We consider each urn in turn. If an urn has received i red balls and j blue balls, we put ij green balls in the urn for the equijoin, or $iI_{j>0}$ green balls for the semijoin;
- The total number of green balls is now the (equi- or semi-) join size.

Urn models were already encountered in the dynamic analysis of tries (see [5]), but there the capacity of the urn varied in time.

2.2. Indicator functions. We introduce the following definitions (κ is an integer-valued function; in the following it counts the number of balls in an urn):

$$\phi_1(\kappa) = I_{\kappa>0}; \qquad \phi_2(\kappa) = \kappa.$$

We shall use upper indices R and B to make precise the relations, or ball colours, we consider. Thus the functions ϕ_2^R and ϕ_2^B are involved in the equijoin, and the functions ϕ_2^R and ϕ_1^B in the

semijoin. We shall also use

$$E_1^i(\phi) := E[\phi(\kappa_1^i)]; \qquad E_{1,2}^{i,j}(\phi) := E[\phi(\kappa_1^i)\phi(\kappa_2^j)]; \qquad Z_1^i := \Pr[\kappa_1^i = 0].$$

2.3. Bi-dimensional processes. Assume that the stochastic processes describing the initial relations R and S are known. There are three ways to combine them, according to the type of correlation allowed on the relations R and S:

(1) The processes may be correlated, in that, at each step, we either update one (and only one) of the relations, or we make a search (query);
(2) The processes may be independent: at each step and for each relation, we can do an insertion, a deletion or a search. In this model, it is possible to update two relations simultaneously, or to query one relation and to update the other one;
(3) We may extend this second model to allow more general probabilities: we define a probability for query, and eight probabilities for updating one or two relations.

For each type of correlation, we can compute the expectation of the join size and the covariance matrix of the bi-dimensional process *(size of R, size of S)*.

3. Birth and death process

We now return to the processes describing the number of balls in an urn, either finite or infinite.

3.1. Infinite urns. The number of balls in one urn is (asymptotically) given by a birth and death process with birth rate and individual death rate given by[1]

$$\lambda(t) = p_{\mathcal{I}}(t)\frac{n}{d}, \qquad \mu(t) = \frac{p_{\mathcal{D}}(t)}{f_1(t)}.$$

The probability that a ball, present at time t_1, survives at time t_2, is

$$ps_{1,2} = \exp\left[-\int_{t_1}^{t_2} \mu(s)\,ds\right].$$

The total number of balls inserted in one urn between times t_1 and t_2, and not deleted at time t_2, follows a Poisson distribution with parameter

$$\rho_{1,2} = \frac{n}{d}\int_{t_1}^{t_2} p_{\mathcal{I}}(u) ps_{1,2}(u, t_2)\,du.$$

3.2. Bounded urns. Such an urn has δ cells; define $\beta := d\delta/n$. At time t_1, the number of balls in any one urn follows (approximately) a binomial distribution with parameters δ and $f_1(t_1)/\beta$.

The next result is Lemma 3 of [2]:

PROPOSITION 1. *Given that we start with k_1 balls in the urn U_i at time t_1, the number of balls $\kappa(t)$ $(t > t_1)$ in the urn U_i is described asymptotically by a birth and death process starting from k_1, with birth rate $\lambda(t) = [\delta - \kappa(t)]f_4(t)$, where $f_4(t) = p_{\mathcal{I}}(t)/[\beta - f_1(t)]$ is the birth rate in a cell, and individual death rate $f_5(t) = p_{\mathcal{D}}(t)/f_1(t)$.*

We can now analyze the distribution of the number of balls in one urn at time t_2, conditioned by the number of balls in the urn at time t_1 (see [2] for details).

[1] We recall that the expectation of the number of balls is asymptotically equal to $nf_1(t)$, with f_1 varying according to the relation scheme.

4. Size of a join

4.1. Theorem. Theorem 6.1 of [3] gives a characterisation of the size of a join as a non-Markovian, Gaussian process of known expectation and covariance:

THEOREM 1. *In the join model, the size $S([nt])$ of the join at time nt is asymptotically given by a non-Markovian Gaussian process with*

$$E[S([nt])] \sim nG(t) \qquad \mathrm{Cov}(S_1, S_2) \sim n\Psi_R(t_1, t_2).$$

The relative error in the density is $O(1/\sqrt{n})$.

4.2. Idea of the proof. The principle of the proof is the same as for the projection. A fundamental result is Lemma 1 of [3], which we recall below.

Define $Y_1 = \sum_{i=1}^{d} \phi(\kappa_1^i)\psi(\lambda_1^i)$: the function κ is relative to red balls, i.e., to the relation R, and the function λ is relative to blue balls, i.e., to the relation S. The index i denotes the number of the urn, and the index 1 shows that we consider the situation at time t_1. We define similarly Y_2 for time t_2.

PROPOSITION 2. *The mean of Y_1 and the covariance of Y_1 and Y_2 are given by*

$$\begin{aligned} E(Y_1) &= dE_1^i(\varphi)E_1^i(\psi); \\ \mathrm{Cov}(Y_1, Y_2) &= dE_{1,2}^{i,i}[\varphi]E_{1,2}^{i,i}[\psi] - dE_{1,2}^{i,j}[\varphi]E_{1,2}^{i,j}[\psi] \\ &\quad + d^2\left[E_1^i[\varphi]E_2^j[\varphi]\ C_{1,2}^{i,j}[\psi] + E_1^i[\psi]E_2^j[\psi]\ C_{1,2}^{i,j}[\varphi] + C_{1,2}^{i,j}[\varphi]C_{1,2}^{i,j}[\psi]\right], \end{aligned}$$

with

$$\begin{aligned} C_{1,2}^{i,j}[\varphi] :&= \sum_{k_1} \Pr(\kappa_1^i = k_1)\varphi(k_1) \sum_{k_2} [\Pr(\kappa_2^j = k_2 | \kappa_1^i = k_1) - \Pr(\kappa_2^j = k_2)]\varphi(k_2) \\ &= E_{1,2}^{i,j}[\varphi] - E_1^j[\varphi]E_2^j[\varphi]. \end{aligned}$$

Proposition 2 of [3] gives the covariance for the "static" structure, when the sizes of the initial relations are assumed known ("non-random" case):

$$E(Y_1) \sim n\ F\left(f_1^R(t_1), f_1^B(t_1)\right); \qquad \mathrm{Cov}(Y_1, Y_2) \sim n\ \psi_{NR}(t_1, t_2).$$

As the number of balls in the urns in no longer known with certainty, but is described by a random variable, this introduces a perturbation on the join size, which can be computed. Lengthy computations give the result.

4.3. Example. Let us take an example to illustrate our results: the equijoin when the urns for both relations are unbounded and when, at each step, either we throw a ball, or we delete a ball, or we make a search. Asymptotically, i.e., for $n \to +\infty$, the expectation of the process *size of equijoin* is

$$E[S([nt])] \sim n\frac{\bar{x}_R\bar{x}_B t^2}{\alpha},$$

and its covariance is

$$\mathrm{Cov}(S_1, S_2) \sim n\left[\frac{\bar{x}_R\bar{x}_B t_1^2 ps_{1,2}^R ps_{1,2}^B}{\alpha} + \frac{\bar{x}_B^2 t_1 t_2}{\alpha^2}\sigma_R^2 t_1 - 2\frac{\bar{x}_B\bar{x}_R t_1 t_2}{\alpha^2}\bar{x}_B\bar{x}_R t_1 + \frac{\bar{x}_R^2 t_1 t_2}{\alpha^2}\sigma_B^2 t_1\right].$$

5. Conclusion

A natural direction for further research would be to implement a toolbox in a computer algebra system such as Maple which, given the structure and dependencies of the initial relations, the kind of relational operation, and the update operations on the database, would compute automatically the moments of the process describing the projection or join size.

Bibliography

[1] Gardy (D.). – Join sizes, urn models and normal limiting distributions. *Theoretical Computer Science, Series A*, vol. 131, n° 2, August 1994.

[2] Gardy (D.) and Louchard (G.). – *Dynamic analysis of some relational data base parameters I: projections.* – Technical Report n° 94-6, Laboratoire Prism, University of Versailles, February 1994.

[3] Gardy (D.) and Louchard (G.). – *Dynamic analysis of some relational data base parameters II: equijoins and semijoins.* – Technical Report n° 94-7, Laboratoire Prism, University of Versailles, February 1994.

[4] Gardy (Danièle). – Sizes of relations: a dynamic analysis. In Salvy (B.) (editor), *Algorithms Seminar 1993-1994*. Institut National de Recherche en Informatique et en Automatique, *Research Report.* – 1994. These proceedings.

[5] Louchard (G.). – Trie size in a dynamic list structure. In Gaudel (M.-C.) and Jouannaud (J.-P.) (editors), *TAPSOFT'93. Lecture Notes in Computer Science*, vol. 668, pp. 719–731. – Springer Verlag, 1993.

Part 5

Miscellany

Elliptic Functions and Modular Forms

François Morain
École Polytechnique

March, 7, 1994

[summary by Daniel Augot]

Abstract

Up to now, there is no good algorithm for computing logarithms in a general finite abelian group. Elliptic curves over finite fields present examples of such groups, and are good candidates for constructing cryptosystems based on exponentiation. To do so, one needs a generator, and to be able to find one, the order of the elliptic curves must be known. It can be computed with machines, and prime numbers up to 250 digits can be dealt with. This first talk introduces the material about elliptic curves, modular forms, ... which is necessary for describing *modular equations*, while the second talk describes algorithms for finding the order of an elliptic curve, specially the "Schoof-Atkin-Elkies" algorithm. Recent work by Couveignes gives an improvement of the method.

1. Elliptic functions

First a whole bunch of definitions, theorems and examples are presented, which are a bit classical.

1.1. Lattices, Eisenstein's series. At the beginning of time, there were lattices:

DEFINITION 1. A *lattice* $\mathbb{L}$ is $\mathbb{L} = \mathbb{Z}\omega_2 + \mathbb{Z}\omega_1$, where $\tau = \omega_1/\omega_2 \in \mathcal{H}$ the upper half-plane. A *cell* is $\{\lambda\omega_1 + \mu\omega_2 + z,\ (\lambda, \omega) \in (0,1)^2\}$, for $z \in \mathbb{C}$.

The following definitions are related to a given lattice.

DEFINITION 2. A meromorphic function f on $\mathbb{C}$ is an *elliptic function* if and only if f is doubly periodic: $\forall z \in \mathbb{C},\ f(z+\omega_1) = f(z+\omega_2) = f(z)$.

PROPOSITION 1. *Let f be an elliptic function. The number of poles and the number of zeroes in a cell is finite. The sums of the residues at poles is* 0. *An elliptic function with no poles is a constant.*

DEFINITION 3 (WEIERSTRASS'S $\wp$ FUNCTION). The *Weierstrass's $\wp$ function* associated to the lattice $\mathbb{L}$ is

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \mathbb{L},\, \omega \neq 0} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right).$$

PROPOSITION 2. *Weierstrass's $\wp$ function is differentiable, and*

$$\wp' = -2 \sum_{\omega \in \mathbb{L}} \frac{1}{(z-\omega)^3}.$$

The functions $\wp$ and $\wp'$ are periodic on $\mathbb{L}$, and the field of elliptic functions is $\mathbb{C}(\wp, \wp')$.

Starting from:

$$\frac{1}{(z-\omega)^2} = \frac{1}{\omega^2\left(1-\frac{z}{\omega}\right)^2} = \frac{1}{\omega^2} + \frac{2z}{\omega^3} + \cdots + \frac{kz^{k-1}}{\omega^{k+1}} + \cdots,$$

the expansion of $\wp$ near the origin is

$$\wp = \frac{1}{z^2} + 2zG_3(\mathbb{L}) + 3z^2G_4(\mathbb{L}) + \cdots + kz^{k-1}G_{k+1}(\mathbb{L}) + \cdots$$

where

$$G_k(\mathbb{L}) = \sum_{\omega\in\mathbb{L},\omega\neq 0} \frac{1}{\omega^2}.$$

We also denote $G_k(\tau)$ the function $G_k(1,\tau)$. The functions $g_2(z)$ and $g_3(z)$ are $g_2(z) = 60G_4(z)$, $g_3(z) = 140G_6(z)$. The *Eisenstein's series* are $E_k(\tau) = G_k(\tau)/(2\zeta(k))$, $k \geq 2$

THEOREM 1. *Let τ be given, and $g_2 = g_2(\tau)$, $g_3 = g_3(\tau)$. Let C be the curve defined by the equation*

$$y^2 = 4x^3 - g_2x - g_3.$$

Then:

(1) *the equation $4x^3 - g_2x^2 - g_3 = 0$ has three distinct roots,*
(2) *the curve C is* parameterized *par $\wp, \wp'$: for each point (x,y) of C, there exists $z \in \mathbb{C}$ such that $(x,y) = (\wp(z), \wp'(z))$.*

Conversely, if C is a curve given by the equation $y^2 = 4x^3 - a_2x - a_3$, such that $4x^3 - a_2x - a_3$ has three distinct roots in $\mathbb{C}$, then there is a lattice $\mathbb{L}$ such that $a_2 = g_2(\mathbb{L})$ and $a_3 = g_3(\mathbb{L})$. The function $\wp_{\mathbb{L}}$ and its derivative yield a parameterisation of C.

2. Modular Functions and modular forms

2.1. Definitions.

DEFINITION 4. The *Poincaré half-plane* is defined as $\mathcal{H} = \{z \in \mathbb{C},\ \mathcal{I}(z) > 0\}$, $\mathcal{I}(z)$ standing for the imaginary part of $z \in \mathbb{C}$. The *modular group* Γ is the set of matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \text{with} \quad (a,b,c,d) \in \mathbb{Z}^4,\ ad - bc = 1.$$

The matrices $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ are the classical generators of Γ. An action of Γ on $\mathcal{H}$ is defined by

$$\forall M \in \Gamma, \forall \tau \in \mathcal{H}, M\tau = \frac{a\tau + b}{c\tau + d}.$$

At last $\mathcal{H}^\star = \mathcal{H} \cup \mathbb{Q} \cup i\infty$ is a compactification of $\mathcal{H}$.

DEFINITION 5. $f : \mathcal{H}^\star \to \hat{\mathbb{C}}$ is a *modular function of weight k* if and only if:

(1) f is meromorphic on $\mathcal{H}$,
(2) $\forall M \in \Gamma, \forall \tau \in \mathcal{H}^\star, f(M\tau) = (c\tau + d)^k f(\tau)$.

If $f(i\infty) \in \mathbb{C}$, then f is a *modular form*, and if $f(i\infty) = 0$, f is a *cusp form*.

EXAMPLE. The Eisenstein's series $E_k(\tau)$ is a modular form of weight k, $k > 2$.

PROPOSITION 3. *There exists no modular form of odd weight k. If f is of weight k and g of weight k', then fg is of weight $k + k'$, f/g of weight $k - k'$.*

EXAMPLE.

- $\Delta(\tau) = (2\pi)^{12}\left(E_4^3(\tau) - E_6(\tau)^2\right)/1728$ is a cusp form of weight 12 ;
- the modular invariant $j(\tau) = 1728 g_2^3(\tau)/\Delta(\tau)$ is a function of weight 0 ;
- $(2\pi)^{-12}\Delta(q)$ can be proven to be equal to $\eta(q)^{24}$, where $\eta(q) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n)$ is the η function.

THEOREM 2. *Let f be a meromorphic function on $\mathcal{H}$. The following statements are equivalent:*

(1) *f is a modular function of weight 0 ;*
(2) *f is the quotient of two modular forms of same weight ;*
(3) *$f \in \mathbb{C}(j)$.*

Let $\mathcal{M}_k$ be the vector space of modular functions of weight k. If $k = 2$ or $k < 0$ then $M_k = \{0\}$, if $k = 4, 6, 8, 10$ then $\mathcal{M}_k$ is of dimension 1 generated by E_k. $\mathcal{M}_1$ is generated by 1.

As a consequence, it is easy to show that $E_8 = E_4^2$, $E_{10} = E_4 E_6$. The Eisenstein series E_2 is not a form, since

$$E_2(-1/\tau) = \tau^2 E_2(\tau) + \frac{12\tau}{2i\pi}.$$

2.2. Modular forms for subgroups. Let Γ_1 be a subgroup of Γ of finite index.

EXAMPLE. Let $\Gamma_0(\ell)$ be the subgroup of matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, with $c \equiv 0 \bmod \ell$. The index $\mu_0(\ell)$ of $\Gamma_0(\ell)$ is $\mu_0(\ell) = \ell\prod_{p|\ell}(1 + 1/p)$.

DEFINITION 6. $\mathcal{F}_1$ is a *fundamental set* for Γ_1 if and only if any point of $\mathcal{H}^\star$ is equivalent (modulo Γ_1) to a unique point in $\mathcal{F}_1$. $\mathcal{F}_1$ is a *fundamental region* if the conditions

$$\tau \in \mathcal{F}_1, \quad \exists M \in \Gamma_1, M \neq 1, M\tau \in \mathcal{F}_1$$

implie that τ belongs to the boundary of $\mathcal{F}_1$.

THEOREM 3. *Let Γ_1 be a finite subgroup of finite index μ, and $\{S_v\}_{1\leq v\leq \mu}$ be a set of coset representatives of Γ_1, i.e.*

$$\Gamma/\Gamma_1 = \{\bar{S_v}\}_{1\leq v\leq \mu}.$$

Then

$$\mathcal{F}_1 = \bigcup_{v=1}^{\mu} S_v(\mathcal{F})$$

is a fundamental region for Γ_1.

2.3. Modular equations.

DEFINITION 7. A function f on $\mathcal{H}^\star$ is a *modular function* for Γ_1 if and only if

(1) f is meromorphic on $\mathcal{H}$,
(2) $\forall M \in \Gamma_1, \forall \tau \in \mathcal{H}^\star, f(M\tau) = f(\tau)$.

It works naturally: if f is a function for a subgroup Γ_1, then $f \circ M$, denoted $f\,|_M$, is a function for the conjugate of Γ_1 by M. A function for a subgroup is a function for its subgroups.

THEOREM 4. *Let f be a function for Γ_1. Set*

$$G(X) = \prod_{v=1}^{\mu} (X - f\,|_{S_v}).$$

The polynomial $G(X)$ can be written

$$G(X) = \sum_{v=0}^{\mu} R_v(j) X^v,$$

where $R_v(j) \in \mathbb{C}(j)$. Then $G(f(q)) = 0$. Such an equation is called a modular equation *for Γ_1. If $f = \sum a_n q^n$ has integer coefficients coefficients, then $G(X, j)$ has integer coefficients.*

EXAMPLE. (Canonical modular equation for $\Gamma_0(\ell)$)
Let $s = 12/\gcd(12, \ell - 1)$, and $v = s(\ell - 1)/12$. The function

$$f(\tau) = \ell^s \left(\frac{\eta(\ell\tau)}{\eta(\tau)} \right)^{2s} = q^v + \sum_{n=v+1}^{\infty} a_n q^n,$$

is a function for $\Gamma_0(\ell)$. The modular equation for f is

$$\Phi_\ell^c(X, j) = (X - f(\tau)) \prod_{k=0}^{\ell-1} (X - f(-1/(1 + k\tau))). \tag{1}$$

Let w_ℓ be the operation associated to

$$\begin{pmatrix} 0 & -1 \\ \ell & 0 \end{pmatrix}.$$

The application w_ℓ is an involution (the *Atkin-Lehmer* involution), and if f is a function for $\Gamma_0(\ell)$, so is $f \circ w_\ell$. Using the Atkin-Lehmer involution, the equation (1) is transformed into

$$P(Y, j) = (Y - \ell^s / f(\tau)) \prod_{k=0}^{\ell-1} (Y - f((\tau + k)/\ell)) = Y^{\ell+1} + \sum_{r=0}^{\ell} C_r(j) Y^r,$$

with $\deg(C_r(j)) \leq v - \frac{rv}{\ell}$. The power-sum symmetric functions

$$S_r = (\ell^s / f(\tau))^r + \sum_{k=0}^{\ell-1} f((\tau + k)/\ell)^r$$

can be computed, and thus the coefficients $C_r(j)$, by Newton's identities.

Bibliography

[1] Atkin (A. O. L.). – The number of points on an elliptic curve modulo a prime. – Preprint, 1988.

[2] Schoeneberg (B.). – *Elliptic modular functions.* – Springer-Verlag, 1974, *Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen*, vol. 203.

[3] Serre (J.-P.). – *Cours d'arithmétique.* – Presses Universitaires de France, 1970.

Implementation of the Schoof-Atkin-Elkies Algorithm

François Morain
École Polytechnique

March, 7, 1994

[summary by Daniel Augot]

After the definitions introduced in the previous talk, algorithms are presented for computing the number points of an elliptic curve on $\mathbb{F}_p$, $p > 3$ a prime number. Three authors have contributed to the problem, Schoof, Atkin and Elkies. The basic ideas are from Schoof and Atkin, Elkies gives a refinement of Schoof's method. A merge of Elkies and Schoof's methods, due to Atkin, is finally adopted.

1. Elliptic Curves

1.1. Group Law. The abelian group structure of an elliptic curve can be used in cryptography, using the difficulty of the discrete logarithm problem. One has to find a generator, and thus must compute the number of points on the elliptic curve over $\mathbb{F}_p$.

An elliptic curve $E(\mathbf{k})$ over a field $\mathbf{k}$ is defined by its (projective) equation

$$y^2z = x^3 + Axz^2 + Bz^3.$$

The group law on $E(\mathbf{k})$ (restricted to the affine plane) is

$$(x_3, y_3) = (x_1, y_1) + (x_2, y_2).$$

where

$$\begin{cases} x_3 &= \lambda^2 - x_1 - x_2 \\ y_3 &= \lambda(x_1 - x_3) - y_1 \end{cases} \quad \text{and } \lambda = \begin{cases} (y_2 - y_1)(x_2 - x_1)^{-1} \text{ if } x_2 \neq x_1, \\ (3x_1^2 + A)(2y_1)^{-1} \text{ otherwise.} \end{cases}$$

The neutral element if $0_E = (0 : 1 : 0)$; the invariant is $j(E) = 2^8 3^3 \frac{A^3}{4A^3+27B^2}$.

THEOREM 1. *$\#E(\mathbb{F}_p) = p + 1 - t$, $|t| \leq 2\sqrt{p}$. $E(\mathbb{F}_p)$ is isomorphic to $E_1 \times E_2$, where $\#E_1 = m_1$, $\#E_2 = m_2$, $m_1 \mid m_2$ and $m_1 \mid p - 1$.*

1.2. Counting points, old. The "Lang-Trotter" method uses the quadratic character:

$$\#E(\mathbb{F}_p) = p + 1 + \sum_{x=0}^{p-1} \left(\frac{x^3 + Ax + B}{p} \right).$$

The cost of computation is $O(p)$, thus convenient for small p only.

Shanks' technique of "Baby Steps, Giant Steps", valid for general groups, can be applied here, at cost $O(p^{\frac{1}{4}})$.

2. (Full) Schoof

2.1. Division polynomials. One can see that the coordinates of nP can be expressed in terms of polynomials in the coordinates of P a point of the elliptic curve:

$$n(X,Y) = \left(\frac{\phi_n(X,Y)}{\psi_n^2(X,Y)}, \frac{\omega_n(X,Y)}{\psi_n^3(X,Y)}\right).$$

DEFINITION 1. The division polynomial $f_n(X,Y)$ is

$$f_n(X,Y) = \begin{cases} \psi_n(X,Y) & \text{for } n \text{ odd,} \\ \psi_n(X,Y)/(2Y) & \text{for } n \text{ even.} \end{cases}$$

The first division polynomials are $f_{-1} = -1$, $f_0 = 0$, $f_1 = 1$, $f_2 = 1$, $f_3(X,Y) = 3X^4 + 6AX^4 + 12BX - A^2$.

The division polynomials can be computed with recurrence relations. They are in fact univariate polynomials, and their degree is of order $\approx n^2/2$. The points $P = (x,y) \in E$ of order ℓ in $E(\overline{\mathbb{F}_p})$ are the points such that $f_\ell(x) = 0$; these points can be described in the algebra $\mathbb{F}_p[X,Y]/\left(Y^2 - (X^3 + AX + B), f_\ell(X)\right)$.

THEOREM 2. *Let $E[\ell] = \{P \in \mathbb{P}_2(\overline{\mathbb{F}_p}), \ell P = 0_E\}$ denotes the set of points of ℓ-division. Then $E[\ell]$ is isomorphic to $\mathbb{Z}/\ell\mathbb{Z} \times \mathbb{Z}/\ell\mathbb{Z}$.*

2.2. The algorithm. Let $\phi : E \to E$ denote the map $(x,y) \mapsto (x^p, y^p)$. It is known that ϕ satisfies the equation $\phi^2 - t\phi + p = 0$, where t is defined in theorem 1.

One works with the equation

$$\left(X^{p^2}, Y^{p^2}\right) + p\,(X,Y) = t\,(X^p, Y^p),$$

in the field $\mathbb{F}_p[X,Y]/\left(Y^2 - (X^3 + AX + B), f_\ell(X)\right)$, finding the value of $t \bmod \ell$ by trial and error.

For different values of ℓ, $t \bmod \ell$ is obtained, and t is found by the Chinese remainder theorem, knowing that $|t| \leq 2\sqrt{p}$.

3. Atkin

The trouble with division polynomials is their important degree, $\approx \ell^2/2$ for points of ℓ-division. Atkin's approach does not use the division polynomials, and is related to the modular equation $\Phi_\ell(j(q^\ell), j(q))$.

3.1. The factorization of a modular equation. The (canonical) modular equation for $j(q^\ell)$ is in strong correspondence with the points of ℓ-division.

It is known from the literature that a modular equation $\Phi_\ell(X, j(E))$ admits one of four types of factorization modulo p:

(1) Φ_ℓ factors into s irreducible polynomials of degree r. Then the equation $X^2 - tX + p \equiv 0 \bmod \ell$ has two roots α and β in $\mathbb{F}_{\ell^2} \setminus \mathbb{F}_\ell$, such that the multiplicative order of α/β is r, and $\ell + 1 = rs$;
(2) Φ_ℓ has a factorization of type $(1)(1)(r)\cdots(r)$. Then the equation $X^2 - tX + p \equiv 0 \bmod \ell$ has two roots α and β in $\mathbb{F}_\ell$, such that the multiplicative order of α/β is r, and $\ell - 1 = rs$;
(3) Φ_ℓ has a factorisation of type $(1)(\ell)$. Then $t^2 \equiv 4p \bmod \ell$;
(4) Φ_ℓ has a factorisation of type $(1)^{\ell+1}$, and $t^2 \equiv 4p \bmod \ell^2$.

3.2. Matching. Having computed the equation Φ_ℓ, the type of the factorization of $\Phi_\ell(X, j(E))$ is found by computing $X^p, X^{p^2}, \ldots$ modulo $\Phi_\ell(X, j(E))$, until the degree of the splitting field of $\Phi_\ell(X, j(E))$ is found.

Then values of t are obtained such that the equation $X^2 - tX + p$ has roots corresponding to types of factorization. Note that, for a given ℓ, many values of $t \bmod \ell$ may be found. This is done for several ℓ_i, $i \in [1, n]$ (say), and then a "matching" is done to get the values of $t \bmod \ell_i$, as follows.

Let $K_i = (\prod \ell_j)/\ell_i$, $i \in [1, n]$. Compute and store the values $(p + 1 - \sum_{i=2}^{n} K_i t_i)P$, and compare with $t_1 K_1 P$ for the values at t_1, P being some point of E. This will match for one set of values $t_i \bmod \ell_i$, then t is recovered by the Chinese remainder theorem.

4. Elkies

The idea of Elkies is to work with polynomials $g_\ell(X)$ with lower degree than the ℓ-division polynomials f_ℓ.

4.1. Subsets of points of ℓ-division. We recall that E is isomorphic to $\mathbb{C}/\mathbb{L}(\omega_1, \omega_2)$. Let us consider $x_r = \wp(r\omega_1/\ell)$, for $1 \leq r \leq d = (\ell - 1)/2$. Let p_1 be the sum $p_1 = x_1 + \cdots + x_d$.

THEOREM 3. *The polynomial $g_\ell(X) = \prod_{r=1}^{d}(X - x_r)$ is in $\mathbb{Q}(A, B, p_1)[X]$.*

The polynomial $g_\ell(X)$ obviously divides the division polynomial $f_\ell(X)$; we have the following tower of extensions: $[\mathbb{Q}(A, B, x_1) : \mathbb{Q}(A, B, p_1)] = d$ and $[\mathbb{Q}(A, B, p_1) : \mathbb{Q}(A, B)] = \ell + 1$.

The polynomial $g_\ell(X)$ describes a subset of $E[\ell]$. The following theorem describes what happens modulo p, and states an interesting case of the factorization of $\Phi_\ell(X, j(E))$.

THEOREM 4. (1) *If $\Phi(X, j(E))$ has a root mod p, then $E[\ell]$ has a cyclic subgroup of order ℓ, whose points have coordinates in $\mathbb{F}_p$. This is equivalent to saying that $t^2 - 4p$ is a square mod ℓ.*

(2) *Then if $t^2 - 4p$ is a square mod ℓ, there is at least one 1-dimensional $\mathbb{F}_\ell$-subspace of $E[\ell]$, which is invariant by ϕ, hence $p_1 \in \mathbb{F}_p$, and $g_\ell(X) \in \mathbb{F}_p[X]$.*

In the hypothesis of theorem 4, $g_\ell(X) \in \mathbb{F}_p[X]$, and we search for an eigenvalue k of ϕ, that is, $k \in [1, \ell]$ such that

$$(X^p, Y^p) = k\,(X, Y) \text{ in } \mathbb{F}_p[X, Y]/\left(Y^2 - (X^3 + AX + B), g_\ell(X)\right).$$

This gives an unique value of $t \bmod \ell$, namely $t \equiv (k^2 + p)/k \bmod \ell$.

4.2. The whole algorithm. The main steps of the algorithm are

(1) Compute a modular equation $\Phi_\ell(F, J)$, where F is a function on $\Gamma_0(\ell)$ (may not be the canonical modular equation).

(2) **If** $\Phi_\ell(F, j(E)) \bmod p$ has at least one root **then** *(Elkies' way)*
   (a) compute a factor $g_\ell(X)$ of $f_\ell(X)$ modulo p, having degree $d = (\ell - 1)/2$.
   (b) fink k such that $(X^p, Y^p) = k(X, Y)$ in $\mathbb{F}_p[X, Y]/\left(Y^2 - (X^3 + AX + B), g_\ell(X)\right)$, and deduce $t \equiv (k^2 + p)/k \bmod \ell$.

   else *(Atkin's way)* $\lambda^2 - t\lambda + p \equiv 0 \bmod \ell$ has two roots α and β such that α/β has order r, r defined by the splitting of $\Phi(F, j(E)) \bmod p$ into r' factors of degree r, with $rr' = \ell + 1$. Eventually do some "matching", and deduce the value of $t \bmod \ell$.

Atkin has described an ingenious way for computing p_1, and the value of $g_\ell(X)$.

5. Conclusion

The implementation is now using fast polynomial arithmetic, which has proven to be superior. The operations includes fast multiplication (with FFT, done by R. Lercier), fast division, gcd, ...

Modular equations for $\ell \leq 500$ can be found, with Chinese remainder theorem on 64 bits. Two curves are studied as examples:

- the INRIA curve: $y^2 = x^3 + 105x + 78153$;
- the POLYTECHNIQUE curve: $y^2 = x^3 + 4589x + 91128$.

Morain's record is POLYTECHNIQUE for a prime of 350 digits, using improvements of Couveignes, for computing values of $t \mod l^n$.

Bibliography

[1] Morain (François). – Calcul du nombre de points sur une courbe elliptique dans un corps fini : aspects algorithmiques. – March 1994. Submitted for publication of the Actes des Journées Arithmétiques 1993.

[2] Schoof (R.). – Elliptic curves over finite fields and the computation of square roots mod p. *Mathematics of Computation*, vol. 44, 1985, pp. 483–494.

[3] Schoof (René). – Counting points on elliptic curves over finite fields. – February 1994. Submitted for publication of the Actes des Journées Arithmétiques 1993.

Piecewise-constant derivative systems and their algorithmic properties

Eugene Asarin
Institute for Information Transmission Problems, Moscow
Université Paris 12

June 6, 1994

[summary by Frédéric Chyzak]

Abstract

E. Asarin deals with simple differential systems: dynamical systems with piecewise-constant derivative—in short, PCD systems. Recently, it has been proved that the reachability problem is decidable in a 2-dimensional space [2]. Asarin and Maler proved in [1] that every Turing machine can be simulated by a 3-dimensional PCD system. Thus, the reachability problem is proved to be undecidable in more than two dimensions. Connections with automata theory and first order logic are also given. This summary is based on [1].

1. PCD systems and transition systems

PCD systems are special cases of differential systems; their solutions are trajectories in a continuous space. Besides, transition systems lead to trajectories in a discrete set. This section recalls results detailed in [1] that prove how transition systems can be simulated by PCD systems.

DEFINITION 1. Let $\mathcal{E}$ be $\mathbb{R}^d$ seen as a d-dimensional Euclidean space and f be a vector field defined from a subset of $\mathcal{E}$ to $\mathcal{E}$. The *dynamical system* on $\mathcal{E}$ with respect to f is the differential system ruled by the equation

$$\frac{d^+x}{dt} = f(x),$$

where $\frac{d^+}{dt}$ is the right derivative and where x is a functional unknown in the time t. A *PCD system* is a dynamical system defined by a piecewise-constant vector field taking a finite number of values. A *trajectory* of a dynamical system is a solution of it.

The definition of a transition system formally looks like the previous one.

DEFINITION 2. Let Q be a set whose elements are called *states* and δ a subset of $Q \times Q$. The *transition system* on Q with respect to δ is the system in the unknown $\sigma : \mathbb{N} \mapsto Q$ satisfying $(\sigma_n, \sigma_{n+1}) \in \delta$ for all $n \in \mathbb{N}$. A transition system is *deterministic* if δ reduces to a function.

All transition systems under consideration in this summary are deterministic though the following results also hold with non deterministic ones.

Let $\xi(t)$ be a solution of a PCD system defined on an interval $[0, r] \subseteq \mathbb{R}^+$. As described in [1], this trajectory can be discretised both in space and time and associated to a state sequence:

(1) first, each subset of $\mathcal{E}$ on which f is constant is associated to a state of Q: this is the discretisation in space;

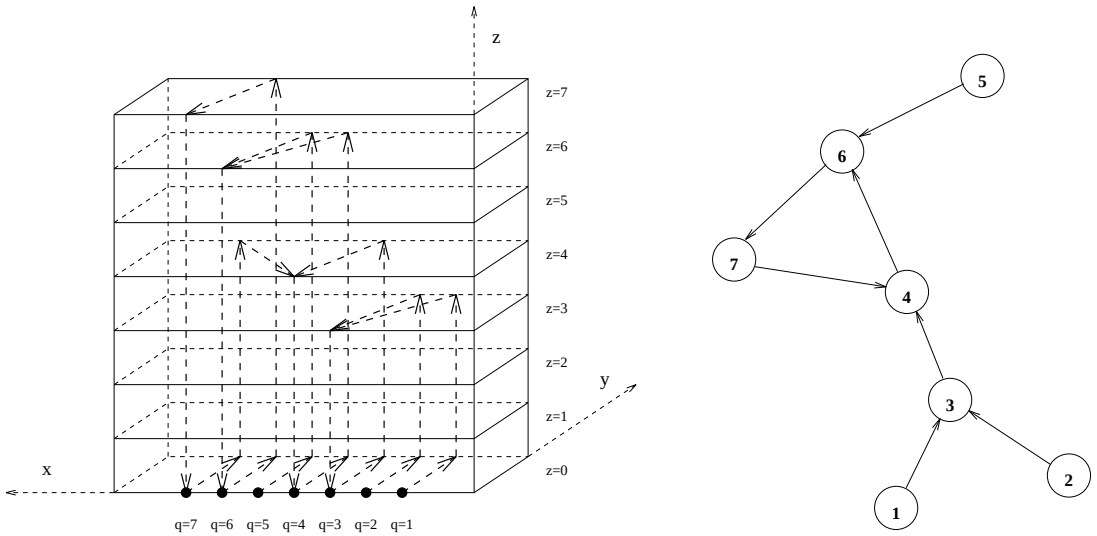

FIGURE 1. A 7-state automaton and its corresponding dynamical system

(2) then, the interval $[0, r]$ on which ξ is defined is divided into successive sub-intervals $[r_n, r_{n+1}]$ on which ξ is constant, so that $\bigcup_{n=0,\ldots,N} [\, r_n, r_{n+1} [\, = [\, 0, r \,[$, where $N \in \mathbb{N} \cup \{\infty\}$: this is the discretisation in time;

(3) finally, the associated state sequence $(\sigma_n)_{n=0,\ldots,N}$ is defined by letting σ_n be the state corresponding to $f([r_n, r_{n+1}])$.

2. Realisation of finite and pushdown automata

In his talk, Asarin focuses on a simple subclass of PCD systems for which the vector field under consideration is constant on *convex polyhedra*: the class of polyhedral-PCD systems. (A convex polyhedron is an intersection of a finite number of indifferently open or closed half-spaces of $\mathcal{E}$.)

The following result proves that PCD systems are at least as powerful as finite automata.

THEOREM 1. *Any finite automaton can be simulated by a 3-dimensional PCD system.*

Rather than a proof, an example of application of this theorem is given in Fig. 1. Note that a state of the automaton is represented by a point of the x-axis, and that the pipes used to lead from one state to another are 1-dimensional. One gets easily convinced that this construction is general.

The results concerning automata with infinite number of states are based on stack machines.

DEFINITION 3. Given an alphabet $\Sigma = \{0, \ldots, k-1\}$, a *stack* S is an element of the set Σ^ω of infinite words on Σ. Such a word is denoted by $s_0 s_1 \cdots$ where $s_n \in \Sigma$ for all $n \in \mathbb{N}$.

For simplicity, only the case $k = 2$ will be dealt with.

The set Σ^ω of possible stacks is certainly not denumerable. Actually, stacks have to be interpreted as binary representations of numbers of $[0, 1]$. Two particular operations can be processed on them.

(1) the PUSH function defined from $\Sigma \times \Sigma^\omega$ to Σ^ω by $\text{PUSH}(s, s_0 s_1 \cdots) = s s_0 s_1 \cdots$;

(2) and the converse POP function defined from Σ^ω to $\Sigma \times \Sigma^\omega$ by $\text{POP}(s_0 s_1 \cdots) = (s_0, s_1 s_2 \cdots)$.

Stack machines can now be defined. As obvious on the following definition, their sets of states $Q \times \Sigma^\omega$ are certainly not denumerable.

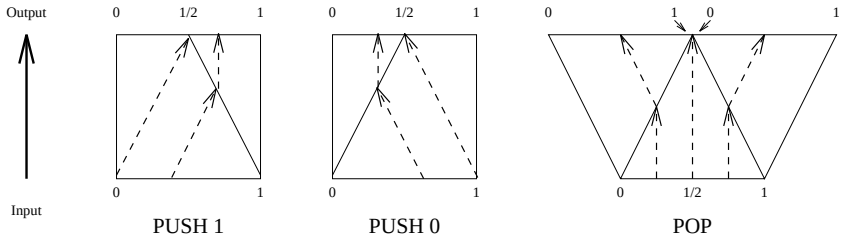

FIGURE 2. PUSH and POP by PCD systems

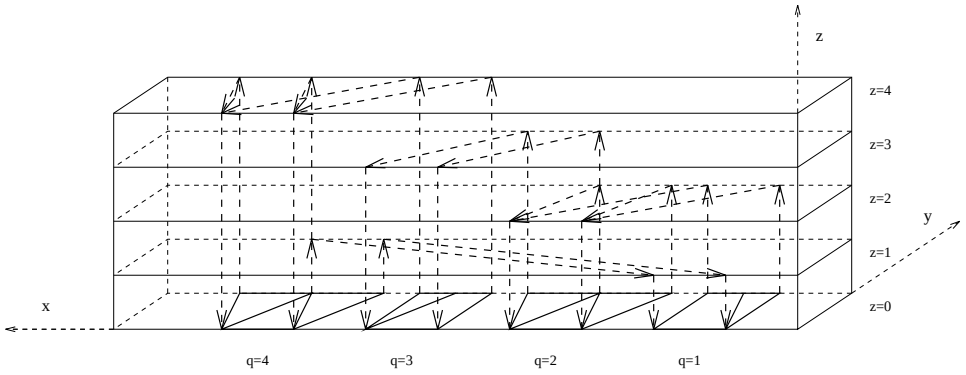

FIGURE 3

DEFINITION 4. A *pushdown automaton* is a transition system $(Q \times \Sigma^\omega, \delta)$ for some finite set $Q = \{q_1, \ldots, q_n\}$ and some δ defined using transitions of either following forms:

(1) the PUSH forms: $\delta(q_i, S) = (q_j, \text{PUSH}(v, S))$, defined for any $v \in \{0, \ldots, k-1\}$;
(2) the POP form: $\delta(q_i, S) = (q_{j_v}, S')$, if $(v, S') = \text{POP}(S)$ and the q_{j_v}'s are elements of Q.

The following result makes use of the continuous nature of $\mathcal{E}$ to prove a result for pushdown automata (PDA) that is similar to the one for finite automata.

THEOREM 2. *Any pushdown automaton can be simulated by a 3-dimensional PCD system.*

Once again, only a sketch of the proof is given. The PUSH and POP transitions are simulated by corresponding "ports" (see Fig. 2).

The pipes are now 2-dimensional strips of $\mathcal{E} = \mathbb{R}^3$, the second dimension encoding a stack. Figure 3 gives the PCD system encoding the PDA defined by:

$$
\begin{aligned}
q_1 &: S := \text{PUSH}(1, S); \text{GOTO } q_2; \\
q_2 &: (v, S) := \text{POP}(S); \text{IF } v = 1 \text{ GOTO } q_2, \text{ELSE GOTO } q_3; \\
q_3 &: S := \text{PUSH}(0, S); \text{GOTO } q_4; \\
q_4 &: (v, S) := \text{POP}(S); \text{IF } v = 1 \text{GOTO } q_1, \text{ELSE GOTO } q_4;
\end{aligned}
$$

3. Reachability problem and realisation of a Turing machine

For a fixed dynamical system $(\mathcal{E}, f)$, given any two points x and x' of $\mathcal{E}$, the reachability problem is to decide the existence of a trajectory ξ and of a $t \in \mathbb{R} - \{0\}$ such that $\xi(0) = x$ and $\xi(t) = x'$.

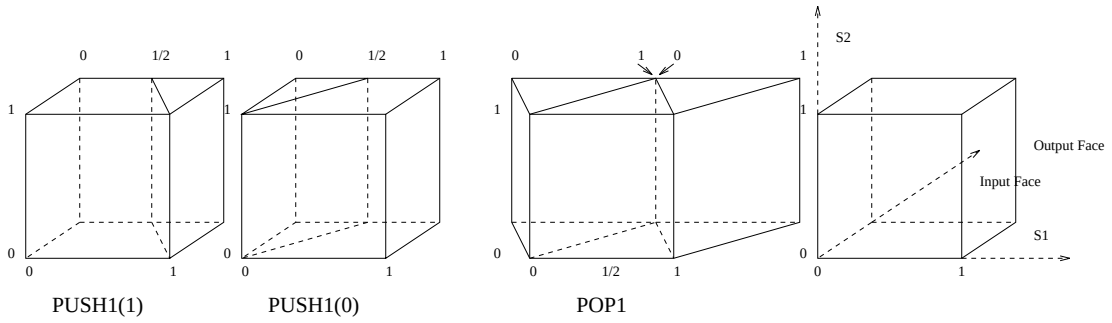

FIGURE 4. PUSH and POP with two stacks

The reachability problem for the class of polyhedral-PCD systems was proved to be decidable in the case of two dimensions by Maler and Pnueli in [2]. The last result about automata can be easily generalised to prove that it becomes undecidable in higher dimensions.

THEOREM 3. *Any pushdown automaton with 2 stacks can be simulated with a 4-dimensional PCD system.*

The idea of the proof is that adding a third dimension to the pipes makes it possible to encode simultaneously two stacks as a point of $[0,1]^2$. Of course, this additional dimension also increments the dimension of $\mathcal{E}$. The "ports" have two be replaced by those of Fig. 4 that push on and pop off the first stack, as well as with corresponding ones for the second stack.

Since any Turing machine can be realised by a pushdown automaton with two stacks, the following corollary holds.

COROLLARY 1. *Every Turing machine can be simulated with a 4-dimensional PCD system.*

Finally, since the halting problem of a Turing machine is equivalent to a reachability problem of a PCD system, the following corollary also holds.

COROLLARY 2. *The reachability problem is undecidable in the case of three (or more) dimensions.*

4. PCD systems and first order logic

Since PCD systems are at least as expressive as Turing machines, any given predicate P on integers $n_1, \ldots, n_k$ can be implemented as a PCD system with the k-tuple $n_1, \ldots, n_k$ in input: the reachability of a special state $q_{\mathtt{true}}$ is equivalent to the provability of $P(n_1, \ldots, n_k)$, the reachability of a special state $q_{\mathtt{false}}$ is equivalent to the provability of $\neg P(n_1, \ldots, n_k)$.

The more general following result will not be proved here.

THEOREM 4. *Any predicate on a tuple of integers with m changes of quantifiers can be computed with a PCD system on a space of dimension $10m$.*

Bibliography

[1] Asarin (Eugene) and Maler (Oded). – On some relations between dynamical systems and transition systems. In *Automata, Languages and Programming. Lecture Notes in Computer Science*, vol. 820. – Springer Verlag, 1994. Proceedings of the 21st International Colloquium, ICALP 94, Jerusalem, Israel.

[2] Maler (O.) and Pnueli (A.). – Reachability analysis of planar multi-linear systems. In Courcoubetis (C.) (editor), *Computer Aided Verification: Proceedings. Lecture Notes in Computer Science*, vol. 697, pp. 194–209. – Springer Verlag, 1993. Proceedings of the 5th International conference, CAV'93, Elounda, Greece.

État de l'art des algorithmes génétiques

Evelyne Lutton
INRIA

28 Mars 1994

[summary by Evelyne Lutton]

Abstract

Les algorithmes génétiques constituent un modèle d'adaptation extrêmement simplifié des systèmes naturels, et sont employés avec succès dans les systèmes artificiels. Ce modèle offre des possibilités non seulement dans le domaine de l'optimisation stochastique, mais aussi dans bien d'autres domaines d'applications, et donne un nouvel éclairage à l'étude des mécanismes de l'évolution naturelle.

Le champ d'application des algorithmes génétiques est très large : il va des applications réelles complexes comme le contrôle du flux de pipelines de gaz, le design de profils d'ailes, ou la planification de trajectoires de robots, à des problèmes plus théoriques de combinatoire, de théorie des jeux, d'économie et d'apprentissage.

1. Introduction

L'idée d'utiliser les principes des processus d'évolution organique en tant que technique d'optimisation globale a émergé indépendamment des deux côtés de l'océan Atlantique il y a une vingtaine d'années. Ces deux approches reposent sur l'imitation du phénomène d'apprentissage collectif d'une population naturelle, basée sur les observations de Darwin et sur la théorie moderne de l'évolution. Ces deux courants ont évolué parallèlement jusqu'à ces dernières années, chacun ayant son champ d'application particulier, tous deux devenant actuellement de plus en plus attirants aussi bien pour les chercheurs que pour les industriels, grâce notamment à la vulgarisation des calculateurs parallèles.

Le courant américain, initialisé par Holland dans les années soixante, a développé ce que l'on appelle les *Algorithmes Génétiques* [6]. Bien qu'ils aient été prévus initialement dans le cadre d'optimisations ou d'adaptations dans le domaine discret, les algorithmes génétiques ont été facilement étendus à l'optimisation sur des domaines continus. En Allemagne, sont apparues à peu près en même temps des méthodes appelées *Stratégies d'Évolution* [5]. Ces méthodes étaient au contraire prévues initialement pour fonctionner sur des domaines continus, et ont été étendues à des applications en optimisation discrète.

Il est relativement malaisé de donner une définition stricte de ce que sont ces approches évolutives (algorithmes génétiques ou stratégie d'évolution) en général. Ce que l'on peut dire, c'est que ces méthodes ont emprunté à la génétique naturelle (et simplifié !) un certain nombre de principes sous-jacents, pour en faire des méthodes algorithmiques, qui s'apparentent dans une certaine mesure aux méthodes d'optimisation combinatoires.

2. Qu'est-ce qu'un algorithme génétique ?

Au premier abord, les algorithmes génétiques peuvent être considérés comme des méthodes d'optimisation stochastique, mais ils ont bien d'autres champs d'applications, comme par exemple en reconnaissance des formes et en intelligence artificielle (systèmes de classeurs ou programmation génétique), en théorie des graphes, en vision et en analyse d'images, en science des matériaux, etc.

L'intérêt d'utiliser un algorithme génétique pour optimiser des fonctions irrégulières est qu'ils savent effectuer une recherche stochastique dans un large espace de recherche, en faisant évoluer un ensemble de solutions (appelée *population*), au lieu d'une seule solution, comme cela est fait classiquement en optimisation stochastique (exemple : un recuit simulé).

Dans l'évolution naturelle, le problème auquel chaque espèce est confrontée est de chercher à s'adapter à un environnement complexe et généralement non statique. Très schématiquement, la "connaissance" acquise par chaque espèce est codée dans les chromosomes de ses membres. Lors des reproductions sexuées, les contenus des chromosomes sont mélangés, modifiés et transmis aux descendants par un certain nombre d'opérateurs génétiques : la mutation, qui se traduit par l'inversion d'une faible partie du matériel génétique, et le croisement (ou recombinaison) qui échange certaines parties des chromosomes des parents. Cette particularité de l'évolution naturelle : la capacité d'une population à explorer son environnement en parallèle et à recombiner les meilleurs individus entre eux, est copiée et exploitée au sein d'un algorithme génétique.

Un résultat formel important mis en évidence par Holland [6] (largement reconnu comme le fondateur du domaine) a été de prouver que, même dans des espaces de recherche larges et complexes, sous certaines conditions, les algorithmes génétiques convergent vers des solutions qui sont à peu près globalement optimales. C'est-à-dire que la population se concentre autour d'un optimum global (théorie des Schémas [4]).

Dans toutes les approches évolutives, les *individus* représentent des solutions ou des points de l'espace de recherche. Cet espace de recherche est appelé *environnement*, c'est sur cet environnement que l'on cherche à maximiser une fonction (positive) appelée *fitness*.

La principale caractéristique des algorithmes génétiques, par rapport aux autres techniques d'inspiration comportementale, est de travailler sur des codages et non sur des solutions réelles. Ces codes sont appelés *chromosomes* et la plupart du temps, en optimisation, ils sont binaires et de longueur fixe. L'algorithme génétique fait donc évoluer sa population de façon à *adapter* les individus à l'environnement, cela se traduit au sens algorithmique du terme par une *maximisation* de la fonction d'évaluation sur les individus de la population.

3. Structure d'un algorithme génétique

La construction d'un algorithme génétique pour une application particulière impose de définir un certain nombre de composantes [2] :

(1) une représentation chromosomale des solutions au problème ;
(2) une méthode de création de la population initiale de solutions ;
(3) une fonction d'évaluation qui joue le rôle de l'environnement, et qui permet d'évaluer les solutions plus ou moins "adaptées" (on parle de *fitness*) ;
(4) des opérateurs génétiques qui modifient la composition des chromosomes des enfants au cours de la reproduction ;
(5) les valeurs des paramètres que l'algorithme génétique emploie (taille de population, probabilités d'application des opérateurs génétiques, etc.).

La structure schématique d'un algorithme génétique est présentée en figure 1. Il existe un grand nombre de variations autour de cette structure.

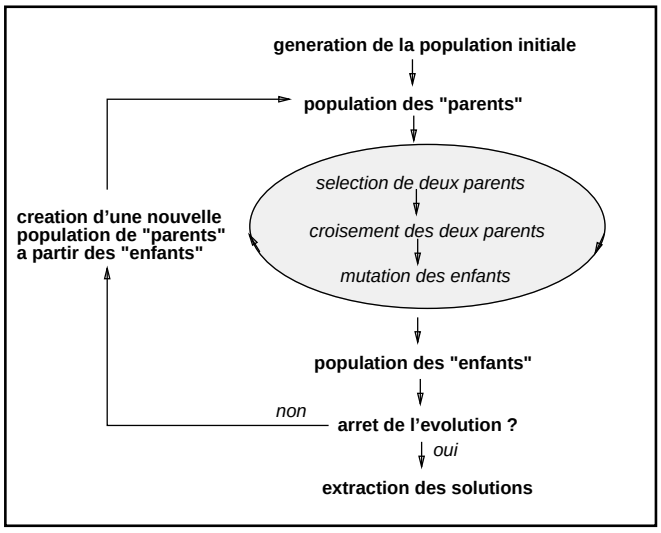

FIGURE 1. Organigramme de l'algorithme génétique simple, cf. [4].

Les solutions (ou *individus*) de la population entrent dans un processus d'évolution. Certaines solutions, meilleures que les autres (leurs fonctions d'évaluations sont meilleures), ont plus de chance de survivre et de transmettre leur patrimoine génétique. La convergence d'un algorithme génétique se traduit par la concentration des individus de la population dans des régions de l'espace de recherche qui présentent un optimum global de la fonction d'évaluation.

La procédure d'initialisation d'un algorithme génétique se fait le plus souvent aléatoirement : les individus de la population initiale sont aléatoirement répartis dans l'espace de recherche. Il est très facile d'introduire des informations *a priori* (sous formes de solutions initiales) dans la population initiale, de façon à accélérer la convergence.

L'évolution de la population se fait par reproduction sexuée. La création d'une nouvelle population (des "enfants") se fait par application de trois opérateurs :

- la *sélection* de deux parents, selon un critère d'adaptation à l'environnement (pour favoriser la reproduction des "bons" individus), c'est là qu'intervient la fonction que l'on cherche à maximiser ;
- le *croisement* des chromosomes des deux parents pour créer les chromosomes de deux enfants (par échange de parties de chromosomes entre les parents) ;
- la *mutation* des chromosomes des enfants (petite perturbation aléatoire sur le chromosome).

La sélection agit comme un opérateur de "concentration", tandis que le croisement a plutôt une action d'exploration. Sélection et croisement favorisent donc la concentration de la population dans des régions de l'espace de recherche où la fonction d'évaluation a des valeurs élevées, avec le risque cependant d'obtenir une convergence prématurée par perte de "diversité génétique" (on dit aussi : par "dérive génétique"). Au contraire la mutation agit comme un opérateur de "dispersion" au sein de la population, et l'on peut démontrer que la mutation permet de maintenir la diversité génétique de la population, et ainsi d'assurer la convergence vers un optimum global. L'efficacité d'un algorithme génétique est donc liée au dosage subtil de ces diverses composantes, afin de converger le plus rapidement possible vers l'optimum global de la fonction à optimiser.

La modélisation et la compréhension des actions de ces trois opérateurs ont été tout d'abord faites au sein de la théorie des Schémas [4, 6]. Plus récemment, les approches par modélisation Markovienne des algorithmes génétiques ont permis de démontrer les effets et l'efficacité de ces

opérateurs [1, 3, 8]. De plus, Davis a proposé il y a peu une démonstration de convergence de l'algorithme génétique simple [3], sur le modèle de la démonstration de la convergence du recuit simulé. Cela lui a permis de proposer une formule de décroissance de la probabilité de mutation (très lente) qui garantit la convergence de l'algorithme vers un optimum global. Le paramètre de probabilité de mutation agit donc de manière similaire à une température dans un recuit simulé.

4. Des applications très nombreuses et très variées

Les algorithmes génétiques classiques, tels que nous venons de les décrire, ont été étendus d'une part par imitation des phénomènes d'évolution naturelle comme la création de niches écologiques (pour l'optimisation de fonctions multimodales) ou la co-évolution de différentes populations (où l'on recherche un état d'équilibre). D'autre part, l'emploi d'autres types de codes que les codes binaires, ont permis des applications orientées Intelligence Artificielle comme les *systèmes de classeurs*, qui manipulent des chromosomes représentant des règles, ou la *programmation génétique*, où les chromosomes représentent directement des programmes arborescents [7].

Les algorithmes génétiques et les algorithmes évolutifs en général intéressent des chercheurs et des ingénieurs de disciplines très diverses, par exemple :

- en optimisation : lorsque les fonctions à optimiser sont complexes, de forte dimensionnalité, irrégulières, mal connues ;
- en intelligence artificielle et sciences cognitives : où l'on exploite plutôt les capacités adaptatives des algorithmes génétiques, et les techniques fondées sur les systèmes de classeurs, (réseaux de neurones, évolution de langages, grammaires) ;
- en robotique : où l'on s'intéresse aux MOBOTS (MObile roBOTS) qui doivent pouvoir se mouvoir et agir dans des environnement inconnus, variables (programmation génétique, systèmes de classeurs) ;
- en physique et en ingénierie : en tant que méthode d'optimisation pour les problèmes réels complexes (pour l'optimisation de structures par exemple) ;
- en économie : pour la modélisation de comportements d'agents par exemple ;
- en traitement d'images, du signal, pour détecter des formes caractéristiques, problème que l'on peut soit comprendre comme une optimisation, soit comme une application de règles de décision (SC) ;
- en théorie des graphes et théorie des jeux : le problème du voyageur de commerce, notamment a beaucoup intéressé les chercheurs.

Bibliographie

[1] Cerf (R.). – Asymptotic convergence of genetic algorithms. – Preprint, 1993.

[2] Davis (L.). – *Genetic Algorithms and Simulated Annealing.* – Pittman, London, 1987.

[3] Davis (T. E.) et Principe (J. C.). – A simulated annealing like convergence theory for the simple genetic algorithm. *In : Proceedings of the Fourth International Conference on Genetic Algorithm*, pp. 174–182. – 1991.

[4] Goldberg (D. A.). – *Genetic Algorithms in Search, Optimization, and Machine Learning.* – Addison-Wesley, janvier 1989.

[5] Hoffmeister (F.) et Baeck (T.). – Genetic algorithms and evolution strategies : Similarities and differences. *In : Parallel Problem Solving from Nature*, pp. 455–470. – 1990.

[6] Holland (J. H.). – *Adaptation in Natural and Artificial System.* – University of Michigan Press, Ann Arbor, 1975.

[7] Koza (J. R.). – *Genetic Programming.* – MIT Press, 1992.

[8] Nix (A. E.) et Vose (M. D.). – Modeling genetic algorithms with Markov chains. *Annals of Mathematics and Artificial Intelligence*, vol. 5, 1992, pp. 79–88.

Contents